

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:07:09 ; Search time 22.5 Seconds
(without alignments)
100.734 Million cell updates/sec

Title: AUDET-909-1

Perfect score: 52

Sequence: 1 eevvpxgmsys 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mmc.*

8: sp_organelle.*

9: sp_phage.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertibrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	73.1	387	16 Q98FX1	Q98fx1 rhizobium l
2	38	73.1	1063	16 Q8RG86	Q8rg86 fusobacteri
3	38	73.1	3472	1 Q74056	Q74056 cenarchaeum
4	37	71.2	840	3 Q9URY8	Q9ury8 schizosacch
5	36	69.2	471	11 Q8R126	Q8r126 mus musculu
6	36	69.2	484	11 Q8VD18	Q8vd18 mus musculu
7	35	67.3	225	10 Q40129	Q40129 lycopersico
8	35	67.3	425	5 Q9XVK4	Q9xvk4 caenorhabdi
9	35	67.3	556	4 Q43733	Q43733 homo sapien
10	35	67.3	583	5 Q9BHA5	Q9bha5 plasmodium
11	35	67.3	583	5 Q9BHA5	Q9bha5 plasmodium
12	35	67.3	670	11 Q01487	Q01487 rattus norv
13	35	67.3	749	16 Q9PDW6	Q9pdm6 xylella fas
14	35	67.3	1902	4 Q14122	Q14122 homo sapien
15	34	65.4	156	3 Q12479	Q12479 saccharomyc
16	34	65.4	219	17 Q971S2	Q971s2 sulfolobus

O28342 archaeoglob
Q96m1 homo sapien
Q92md6 rhizobium m
Q98bp5 rhizobium l
Q9pd2 ureaplasma
Q9ur4 penicillium
Q8xt05 ralstonia s
Q96914 homo sapien
Q8tx62 methanopyru
Q9cxq4 mus musculu
O28330 archaeoglob
Q8vua8 lactococcus
Q8re56 fusobacteri
P74187 synechocyst
Q9m3c0 arabidopsis
Q9g9q9 soil-borne
Q9g9q5 soil-borne
Q9g9q7 soil-borne
Q9d1n1 soil-borne
Q9d1n1 soil-borne
Q9g9q7 soil-borne
Q9g9q7 soil-borne
Q9fml4 arabidopsis
Q9x0u3 thermotoga
Q9v914 drosophila
Q96cs0 homo sapien
Q9d2x9 mus musculu
Q8r8k6 thermoanser
Q96mb2 homo sapien

ALIGNMENTS

RESULT 1

Q98FX1 ID Q98FX1 PRELIMINARY; PRT; 387 AA.
AC Q98FX1; 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE Hippurate hydrolase.
GN MUR3583..
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti";
RL DNA Res 7:331-338(2000).
EMBL; AP003002; BAB50445.1; --
DR InterPro; IPR002933; Peptidase_M20.
DR Pfam; PF01546; Peptidase M20; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 387 AA; 41180 MW; 131BFF8E64306829 CRC64;

Query Match 73.1%; Score 38; DB 16; Length 387;

Best Local Similarity 60.0%; Pred. No. 9.8;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10

Db 367 DEAIHPGMSY 376

```
RESULT 2
Q8RG86 PRELIMINARY; PRT; 1063 AA.
ID Q8RG86
AC Q8RG86;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5).
GN FN0422.
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=1189109;
RA Kapural V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fongstein M., Kyriakides N., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586."
RL J. Bacteriol. 184:2005-2018(2002).
DR EMBL; AF010554; AAL94625.1; -
KW Ligase; Complete proteome.
SQ SEQUENCE 1063 AA; 118008 MW; 39700E10B7CCE411 CRC64;

Query Match 73.1%; Score 38; DB 16; Length 1063;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EVVPXGMSYS 11
|:|:|:|:|
Db 195 EIVPGLNYS 204

RESULT 3
O74056 PRELIMINARY; PRT; 3472 AA.
ID O74056
AC O74056;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 367.1 kDa protein.
OS Cenarchaeum symbiosum.
OC Archaea; Crenarchaeota; Cenarchaeum.
OX NCBI_TaxID=45770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B;
RX MEDLINE=98422450; PubMed=9748430;
RA Schleper C., DeLong E.F., Preston C.M., Feldman R.A., Wu K.Y.,
RA Swanson R.V.;
RT "Genomic analysis reveals chromosomal variation in natural populations
RT of the uncultured psychrophilic archaeon Cenarchaeum symbiosum."
RL J. Bacteriol. 180:5003-5009(1998).
CC -1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL; AF083072; AAC62699.1; -
DR InterPro; IPR000515; BPD transp.
DR Pfam; PF001680; WD40.
DR SMART; SM00320; WD40; 4
DR PROSITE; PS00402; BPD_TRANSP_INN_MEMBER; UNKNOWN_1.
KW Hypothetical protein; Repeat; WD repeat.
SQ SEQUENCE 3472 AA; 367058 MW; 37F80707030F9355 CRC64;

Query Match 73.1%; Score 38; DB 1; Length 3472;
Best Local Similarity 54.5%; Pred. No. 1e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 BEVVPXGMSYS 11
```

```
Db 2294 EDVIPRGISPS 2304

|:|:|:|:|:|
RESULT 4
Q9URY8 PRELIMINARY; PRT; 840 AA.
ID Q9URY8
AC Q9URY8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Probable sulfate permease.
GN SPAC869.05C.
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RA Hunt C., Aves S., McDougall R.C., Rajandream M.A., Barrell B.G.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL132779; CAB60015.1; -
DR InterPro; IPR002645; STAS.
DR InterPro; IPR001902; Sulfate_transp.
DR Pfam; PF01740; STAS; 1.
DR Pfam; PF00916; Sulfate_transp; 1.
DR TIGRFAMs; TIGR00815; sulP; 1.
SQ SEQUENCE 840 AA; 93517 MW; ED4833E162B69077 CRC64;

Query Match 71.2%; Score 37; DB 3; Length 840;
Best Local Similarity 77.8%; Pred. No. 37;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VVPXGMSYS 11
|:|:|:|:|
Db 135 VVPOGMSYA 143

RESULT 5
Q8R126 PRELIMINARY; PRT; 471 AA.
ID Q8R126
AC Q8R126;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 54.5 kDa protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025810; AAH25810.1; -
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 471 AA; 54506 MW; E0DA685C374A9760 CRC64;

Query Match 69.2%; Score 36; DB 11; Length 471;
Best Local Similarity 60.0%; Pred. No. 32;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 EVVPXGMSYS 11
|:|:|:|:|
Db 226 EVIPAGASYN 235

RESULT 6
Q8VD18
```

```

ID Q8VD18 PRELIMINARY; PRT; 484 AA.
AC Q8VD18;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to glioma tumor suppressor candidate region gene 2.
GN AW536441.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SALIVARY GLAND;
RA Strausberg R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC017637; AAH17637.1; -.
DR MGD; MGI:2138595; AW536441.
SQ SEQUENCE 484 AA; 55835 MW; BBB45F3B4BE02A36 CRC64;

Query Match 69.2%; Score 36; DB 11; Length 484;
Best Local Similarity 60.0%; Pred. No. 33;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11
|::|::|
DB 239 EVIPAGASYN 248

RESULT 7
Q40129 PRELIMINARY; PRT; 225 AA.
AC Q40129;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical 25.2 kDa protein precursor.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF36; TISSUE=PISTIL;
RX MEDLINE=95375233; PubMed=7647301;
RA Milligan S.B.; Gasser C.S.;
RT "Nature and regulation of pistil-expressed genes in tomato.";
RL Plant Mol. Biol. 28:691-711(1995).
DR EMBL; U20592; AAA80497.1; -.
DR InterPro; IPR002160; Kunitz legume.
DR Pfam; PF00197; Kunitz_legume; 1.
DR ProDom; PD000891; Kunitz_legume; 1.
DR SMART; SM00452; STI; 1.
DR PROSITE; PS00283; SOYBEAN KUNITZ; UNKNOWN_1.
KW Hypothetical protein; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 225 UNKNOWN.
SQ SEQUENCE 225 AA; 25188 MW; 1074C261D20CFDAD CRC64;

Query Match 67.3%; Score 35; DB 10; Length 225;
Best Local Similarity 54.5%; Pred. No. 24;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVVPXGMSYS 11
:|||||::|
DB 32 DEVVNGKTYA 42

RESULT 8
Q9XVK4 PRELIMINARY; PRT; 425 AA.
ID Q9XVK4
AC Q9XVK4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

```

```

DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE R10D12.10 protein.
GN R10D12.10
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Percy C.M.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z81109; CAB03241.1; -.
DR InterPro; IPR000719; Euk_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Transferase.
SQ SEQUENCE 425 AA; 49410 MW; 5D96E29B08C8E9D6 CRC64;

Query Match 67.3%; Score 35; DB 5; Length 425;
Best Local Similarity 50.0%; Pred. No. 48;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVVPXGMSYS 10
|::|::|
DB 335 EQIVPGGLQY 344

RESULT 9
Q43733 PRELIMINARY; PRT; 556 AA.
ID Q43733
AC Q43733;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-1998 (TrEMBLrel. 06, Last annotation update)
DE DNA binding protein (fragment).
GN DJ451B15.2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tubby B.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z98050; CAB10847.1; -.
FT NON_TER 1
SQ SEQUENCE 556 AA; 59059 MW; EC800E403FB2528 CRC64;

Query Match 67.3%; Score 35; DB 4; Length 556;
Best Local Similarity 66.7%; Pred. No. 64;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11
|||::|
DB 244 VVPAGLTYS 252

RESULT 10
Q9BHA5 PRELIMINARY; PRT; 583 AA.
ID Q9BHA5
AC Q9BHA5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

```

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Choline transporter.
 GN SCT1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OK NCBI_TaxID=5833;
 RP [1]
 RN SEQUENCE FROM N.A.
 RA Ben Mamoun C., Gluzman I.V., Goldberg D.E.;
 RT "Plasmodium falciparum choline transporter (pfSCT1) gene";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY007372; AAK14816.1; -.
 DR EMBL; AY007375; AAG17947.1; -.
 DR InterPro; IPR002123; Acyltransferase.
 DR Pfam; PF01553; Acyltransferase; 1.
 SQ SEQUENCE 583 AA; 66917 MW; 2B2BFAE3E395E049 CRC64;

Query Match 67.3%; Score 35; DB 5; Length 583;
 Best Local Similarity 55.6%; Pred. No. 67;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VVPXGMSYS 11
 : : | : | : |

Db 227 IIPVGLSYS 235
 : : | : | : |

RESULT 11

ID Q9BH83 PRELIMINARY; PRT; 583 AA.
 AC Q9BH83;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Choline transporter.
 GN SCT1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OK NCBI_TaxID=5833;
 RN [1]
 RA Ben Mamoun C., Gluzman I.V., Goldberg D.E.;
 RT "Plasmodium falciparum choline transporter (pfSCT1) gene";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY007374; AAK14818.1; -.
 DR EMBL; AY007373; AAK14817.1; -.
 DR InterPro; IPR002123; Acyltransferase.
 DR Pfam; PF01553; Acyltransferase; 1.
 SQ SEQUENCE 583 AA; 66918 MW; 4DF83D7530527474 CRC64;

Query Match 67.3%; Score 35; DB 5; Length 583;
 Best Local Similarity 55.6%; Pred. No. 67;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VVPXGMSYS 11
 : : | : | : |

Db 227 IIPVGLSYS 235
 : : | : | : |

RESULT 12

ID Q01487 PRELIMINARY; PRT; 670 AA.
 AC Q01487;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE DNA-binding protein AP-BP2 (ALPHA1-antitrypsin promoter binding protein 2) (Fragment).
 GN Rattus norvegicus (Rat).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OK NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=SPRAGUE-DAWLEY; TISSUE=THYROID;
 RX MEDLINE=91187610; PubMed=1901405;
 RA Mitchemore C., Traboni C., Cortese R.;
 RT "Isolation of two cDNAs encoding zinc finger proteins which bind to
 the alpha 1-antitrypsin promoter and to the major histocompatibility
 complex class I enhancer";
 RL Nucleic Acids Res. 19:141-147 (1991).
 CC -!- FUNCTION: BINDS TO THE ALPHA1-ANTITRYPSIN PROMOTER. TO THE KAPPA
 CC IMMUNOGLOBULIN GENE ENHANCER, AND TO THE MAJOR HISTOCOMPATIBILITY
 CC COMPLEX CLASS I ENHANCER; IT PLAYS A ROLE AS TRANSCRIPTIONAL
 CC REGULATOR. CONCERNING THE ALPHA1-ANTITRYPSIN EXPRESSION IT MAY ACT
 CC AS A NEGATIVE REGULATOR BY INTERFERING AT THE ALPHA1-ANTITRYPSIN
 CC PROMOTOR WITH ANOTHER NUCLEAR PROTEIN, CALLED LF-B.
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.
 CC -!- MISCELLANEOUS: ACIDIC AMINO ACID REGION MAY BE INVOLVED IN
 CC TRANSCRIPTIONAL ACTIVATION.
 CC -!- SIMILARITY: HIGHLY SIMILAR TO AT-BP1 AND TO THE PRDII-BF1 FACTOR.
 DR EMBL; X54250; CAA38151.1; -.
 DR HSSP; P15822; 1BBO.
 DR InterPro; IPR000822; Znf_C2H2.
 DR Pfam; PF00096; Zf-C2H2; 2.
 DR SMART; SM00355; Znf_C2H2; 2.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 2.
 KW Zinc-finger; Nuclear protein; DNA-binding; Transcription regulation;
 KW Metal-binding; Multigene family.
 FT NON TER 1
 FT DOMAIN 54 104 ZINC-FINGERS.
 FT DOMAIN 140 160 ACIDIC.
 FT ZN FING 54 74 C(2)H(2) CLASS.
 FT ZN FING 82 104 C(2)H(2) CLASS.
 SQ SEQUENCE 670 AA; 71233 MW; CDD2324152590C17 CRC64;

Query Match 67.3%; Score 35; DB 11; Length 670;
 Best Local Similarity 66.7%; Pred. No. 78;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VVPXGMSYS 11
 | | | : | : |

Db 376 VVPAGLYTS 384
 | | | : | : |

RESULT 13

ID Q9PDM6 PRELIMINARY; PRT; 749 AA.
 AC Q9PDM6;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Topoisomerase IV subunit.
 GN XFI353.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xylella.
 OK NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9A5C;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
 RA Facinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
 RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuranae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,

RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Fesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silva J.F., Silvestri M.L.Z., Silvestri W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tesi S.M., Teuhako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen *Xylella fastidiosa*."
RL Nature 406:151-159(2000).
DR EMBL: AE003967; AAF84162.1; -;
DR HSSP: P09097; IAB4.
DR InterPro: IPR002205; DNA_topoisomIV.
DR Pfam: PF00521; DNA_topoisomIV; 1.
DR ProDom: PD000742; DNA_topoisomIV; 1.
DR SMART: SM00434; TOP4C; 1.
DR TIGRFAMs: TIGR01062; parC_Gneg; 1.
KW Complete proteome.
SQ SEQUENCE 749 AA; 83344 MW; 4DCD10F480E0257 CRC64;

Query Match 67.3%; Score 35; DB 16; Length 749;
Best Local Similarity 77.8%; Pred. No. 88;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPGMSY 10
DB 526 EVDPSGMSY 534

RESULT 14
Q14122 PRELIMINARY; PRT; 1902 AA.
AC Q14122;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
DE DNA-binding protein (Mbp-1) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90205817; PubMed=2108316;
RA Baldwin A.S., LeClair K.P., Singh H., Sharp P.A.;
RT "A large protein containing zinc finger domains binds to related
RT sequence elements in the enhancers of the class I major
RT histocompatibility complex and kappa immunoglobulin genes."
RL Mol. Cell. Biol. 10:1406-1414(1990).
DR EMBL: M32019; AAA17534.1; -;
DR HSSP: P15822; IBB0.
DR InterPro: IPR000822; Znf_C2H2.
DR Pfam: PF00096; zf-C2H2; 3.
DR SMART: SM00355; Znf_C2H2; 2.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 2.
DR PROSITE: PS0157; ZINC_FINGER_C2H2_2; 2.
KW DNA-binding; Metal-binding; Zinc-finger.
FT NON TER 1
SQ SEQUENCE 1902 AA; 207457 MW; 35BBF0D961B6ED20 CRC64;

Query Match 67.3%; Score 35; DB 4; Length 1902;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

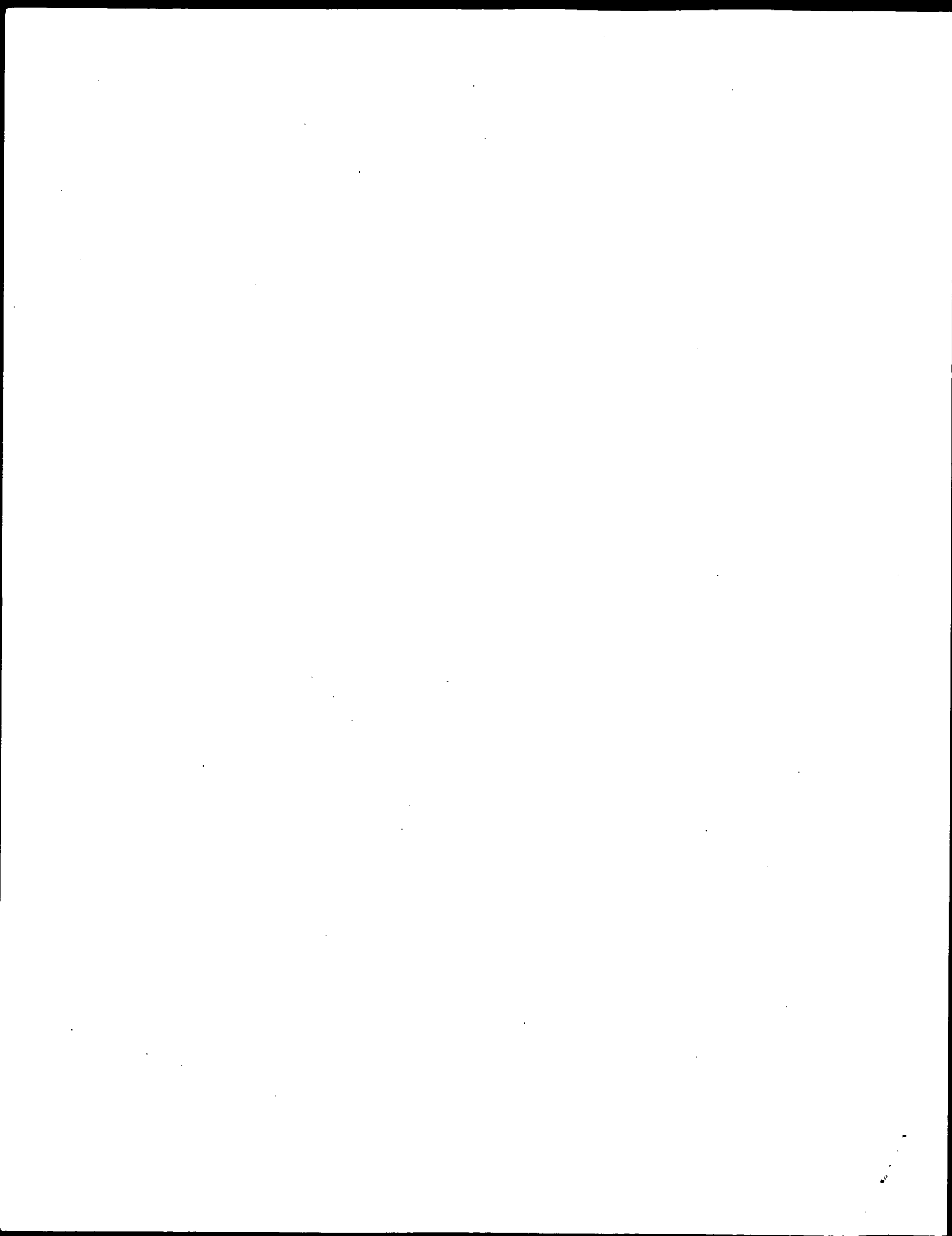
QY 3 VVPGXMSYS 11
DB 1590 VVPAGLTYS 1598

RESULT 15
Q12479 PRELIMINARY; PRT; 156 AA.
AC Q12479;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)
DE ORF YOR013W.
GN YOR013W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA De haan M., Grivell L.A., Maarse A.C.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MIPS;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=FY1679;
RA De haan M., Maarse A.C., Grivell L.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=FY1679;
RX MEDLINE=94019318; PubMed=8413243;
RA Dumont M.E., Schlichter J.B., Cardillo T.S., Hayes M.K., Bethlenny G.,
RA Sherman F.;
RT "CYC2 encodes a factor involved in mitochondrial import of yeast
RT cytochrome c."
RL Mol. Cell. Biol. 13:6442-6451(1993).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=FY1679;
RX MEDLINE=94169519; PubMed=7764548;
RA Lee Y.S., Shimizu J., Yoda K., Yamasaki M.;
RT "Molecular cloning of a gene, DHS1, which complements a drug-
RT hypersensitive mutation of the yeast *Saccharomyces cerevisiae*."
RL Biosci. Biotechnol. Biochem. 58:391-395(1994).
DR EMBL: Z74920; CAA99201.1; -;
DR EMBL: X87331; CAA60762.1; -;
DR SGD: S0005539; YOR013W.
SQ SEQUENCE 156 AA; 17881 MW; 380442B74C272B41 CRC64;

Query Match 65.4%; Score 34; DB 3; Length 156;
Best Local Similarity 66.7%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPGXMSY 10
DB 50 EVNPLGNDY 58

Search completed: June 4, 2003, 13:13:23
Job time : 24.5 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:04:34 ; Search time 6.25 Seconds
(without alignments)
72.998 Million cell updates/sec

Title: AUDET-909-1

Perfect score: 52

Sequence: 1 eevvpxgmsys 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	38	73.1	1058	1 CARB_FUSNN	Q8rg86 fusobacteri
2	37	71.2	877	1 SULH_SCHPO	O74377 schizosacch
3	36	69.2	102	1 Y1IK_TYDVA	P31619 tobacco yel
4	36	69.2	1498	1 Y1A9_CLOAB	Q04351 clostridium
5	35	67.3	2717	1 ZEP1_HUMAN	P15822 homo sapien
6	34	65.4	788	1 CY14_NEUCR	P23622 neurospora
7	34	65.4	1499	1 A10C_HUMAN	O60312 homo sapien
8	33	63.5	116	1 RL20_MYCPU	Q98qv0 mycoplasma
9	33	63.5	253	1 Y990_CAMJE	P45489 campylobact
10	33	63.5	426	1 AROA_VIBCH	Q9krb0 vibrio chol
11	33	63.5	466	1 CC37_SCHPO	O94740 schizosacch
12	33	63.5	478	1 GSR2_HUMAN	Q9nzm5 homo sapien
13	33	63.5	890	1 BCN5_CLOPE	P08696 clostridium
14	33	63.5	1394	1 LTBS_HUMAN	P22064 homo sapien
15	33	63.5	1401	1 RPOC_VIBCH	Q14766 homo sapien
16	33	63.5	1595	1 LTBL_HUMAN	Q00918 rattus norv
17	33	63.5	1712	1 LTB1_RAT	Q04827 rattus norv
18	32	61.5	288	1 CGD2_RAT	P30279 homo sapien
19	32	61.5	289	1 CGD2_HUMAN	P30280 mus musculu
20	32	61.5	289	1 CGD2_MOUSE	Q90459 brachydanio
21	32	61.5	291	1 CGD1_BRARE	P50755 xenopus lae
22	32	61.5	291	1 CGD1_XENLA	P49706 gallus gall
23	32	61.5	291	1 CGD2_CHICK	P53782 xenopus lae
24	32	61.5	291	1 CGD2_XENLA	P55169 gallus gall
25	32	61.5	292	1 CGD1_CHICK	P30281 homo sapien
26	32	61.5	292	1 CGD3_HUMAN	P24385 mus musculu
27	32	61.5	295	1 CGD1_HUMAN	P25322 mus musculu
28	32	61.5	295	1 CGD1_MOUSE	P39948 rattus norv
29	32	61.5	295	1 CGD1_RAT	P44677 haemophilus
30	32	61.5	427	1 TOLB_HAETN	O75355 homo sapien
31	32	61.5	529	1 ENP3_HUMAN	Q9y616 homo sapien
32	32	61.5	591	1 OAT6_HUMAN	P52384 human herpe
33	32	61.5	726	1 PRTP_HSV6U	

RESULT 1

ID	CARB_FUSNN	STANDARD;	PRT;	1058 AA.
AC	Q8RG86;			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-phosphate synthetase ammonia chain).			
GN	CARB OR FN0422.			
OS	Fusobacterium nucleatum (subsp. nucleatum).			
OC	Bacteria; Fusobacteria; Fusobacterium.			
OX	NCBI_TaxID=76856;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 25586;			
RX	MEDLINE=21886334; PubMed=11889109;			
RA	Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A., Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L., Vasileva O., Chu L., Kogan Y., Chaga O., Goltzman E., Hernal A., Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R., Fonstein M., Kyripides N., Overbeek R.;			
RT	"Genome sequence and analysis of the oral bacterium Fusobacterium nucleatum strain ATCC 25586";			
RL	J. Bacteriol. 184:2005-2018(2002).			
CC	-I- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP + phosphate + L-glutamate + carbamoyl phosphate.			
CC	-I- COFACTOR: Binds three manganese ions (By similarity).			
CC	-I- PATHWAY: Arginine biosynthesis.			
CC	-I- PATHWAY: Pyrimidine biosynthesis; first step.			
CC	-I- SUBUNIT: Composed of two chains; the small (or glutamine) chain promotes the hydrolysis of glutamine to ammonia, which is used by the large (or ammonia) chain to synthesize carbamoyl phosphate (By similarity).			
CC	-I- SIMILARITY: BELONGS TO THE CARB FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; AE010554; AAL94625.1; ALT_INIT.			
DR	InterPro; IPR005483; CPase_L.			
DR	InterPro; IPR005479; CPase_L_D2.			
DR	InterPro; IPR005480; CPase_L_D3.			
DR	InterPro; IPR005481; CPase_L_N.			
DR	InterPro; IPR004362; MGS_Like.			
DR	Pfam; PF020289; CPasease_L_Chain; 2.			
DR	Pfam; PF02786; CPasease_L_D2; 2.			
DR	Pfam; PF02787; CPasease_L_D3; 1.			
DR	Pfam; PF02142; MGS; 1.			
DR	PRINTS; PR00098; CPSASE.			
DR	PROSITE; PS00866; CPSASE_1; 2.			

P32784 saccharomyc
Q62671 rattus norv
P28931 tomato aspe
P16916 escherichia
P16918 escherichia
P16917 escherichia
Q95071 homo sapien
P17280 chimpanzee
Q10676 mycobacteri
O59619 pyrococcus
Q9nnu2 halobacteri
O05756 mycobacteri

ALIGNMENTS

```

DR PROSITE; PS00867; CPSASE 2; 2.
KW Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;
KW ATP-binding; Manganese; Complete proteome.
FT DOMAIN 1 401 CARBOXYPHOSPHATE SYNTHETIC DOMAIN.
FT DOMAIN 402 546 OLIGOMERIZATION DOMAIN.
FT DOMAIN 547 929 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.
FT DOMAIN 930 1058 ALLOSTERIC DOMAIN.
FT REPEAT 1 546
FT REPEAT 547 1058
FT NP_BIND 153 210 ATP (POTENTIAL).
FT NP_BIND 302 352 ATP (POTENTIAL).
FT METAL 284 284 MANGANESE 1 (BY SIMILARITY).
FT METAL 298 298 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 300 300 MANGANESE 2 (BY SIMILARITY).
FT METAL 820 820 MANGANESE 3 (BY SIMILARITY).
FT METAL 832 832 MANGANESE 3 (BY SIMILARITY).
SQ SEQUENCE 1058 AA; 117451 MW; ED7037AF77C1E39F CRC64;

Query Match 73.1%; Score 38; DB 1; Length 1058;
Best Local Similarity 60.0%; Pred. No. 6.2;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPGXGMSYS 11
|:|:|:|:|
Db 190 EIVPGLNYS 199

RESULT 2
SULH_SCHPO STANDARD; PRT; 877 AA.
AC O74377;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable sulfate permease C3H7.02.
GN SPEC3H7.02.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skellern J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Furnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg J.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: HIGH AFFINITY UPTAKE OF SULFATE INTO THE CELL (BY
CC SIMILARITY).

```

```

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE SLC26A FAMILY OF TRANSPORTERS.
CC -!- SIMILARITY: CONTAINS 1 STAS DOMAIN.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL031261; CAA20298.1; -.
CC InterPro; IPR002645; STAS.
CC InterPro; IPR001902; Sulfate transp.
CC Pfam; PF00916; Sulfate_transp; 1.
CC TIGR; PF01740; STAS; 1.
CC TIGR; TIGR00815; sulp; 1.
CC PROSITE; PS01130; SLC26A; 1.
CC PROSITE; PS0801; STAS; 1.
KW Transport; Transmembrane.
FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 161 181 POTENTIAL.
FT TRANSMEM 186 206 POTENTIAL.
FT TRANSMEM 221 241 POTENTIAL.
FT TRANSMEM 243 263 POTENTIAL.
FT TRANSMEM 292 312 POTENTIAL.
FT TRANSMEM 329 349 POTENTIAL.
FT TRANSMEM 384 404 POTENTIAL.
FT TRANSMEM 424 444 POTENTIAL.
FT TRANSMEM 461 481 POTENTIAL.
FT TRANSMEM 484 504 POTENTIAL.
FT TRANSMEM 518 538 POTENTIAL.
FT TRANSMEM 543 563 POTENTIAL.
FT DOMAIN 594 747 STAS.
SQ SEQUENCE 877 AA; 96373 MW; 56995A8493371B43 CRC64;

Query Match 71.2%; Score 37; DB 1; Length 877;
Best Local Similarity 77.8%; Pred. No. 8.3;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPGXGMSYS 11
|:|:|:|:|
Db 148 VVPGQMSYA 156

RESULT 3
YIIK_YIYVA STANDARD; PRT; 102 AA.
AC P31619;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-OCT-1993 (Rel. 27, Last annotation update)
DE Hypothetical 11.2 kDa protein.
OS Tobacco yellow dwarf virus (strain Australia) (TYDV).
OC Viruses; ssDNA viruses; Geminiviridae; Mastrevirus.
OX NCBI_TaxID=31599;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92188538; PubMed=1546458;
RA Morris B.A.M., Richardson K.A., Haley A., Zhan X., Thomas J.E.;
RT "The nucleotide sequence of the infectious cloned DNA component of
RT tobacco yellow dwarf virus reveals features of geminiviruses
RT infecting monocotyledonous plants.";
RL Virology 187:633-642(1992).
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

CC or send an email to license@isb-sib.ch).

CC EMBL; M81103; AAA47947.1; -
 DR PIR; A42452; A42452.
 DR InterPro; IPR002621; Gemini mov.
 DR Pfam; PF01708; Gemini_mov; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 102 AA; 11178 MW; A40ECF1E0AF55B67 CRC64;

Query Match 69.2%; Score 36; DB 1; Length 102;
 Best Local Similarity 60.0%; Pred. No. 1.4;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11
 :||| :|||
 Db 7 QVVPSPGINYS 16

RESULT 4
 YIA9_CLOAB STANDARD; PRT; 1498 AA.
 ID YIA9_CLOAB
 AC Q04351;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein CAC3709.
 GN CAC3709.
 OS Clostridium acetobutylicum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE=21359325; PubMed=11466286;
 RA Noelling J., Breton G., Omeichenko M.V., Makarova K.S., Zeng Q.,
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
 RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
 RA Bennett G.N., Koonin E.V., Smith D.R.;
 RT "Genome sequence and comparative analysis of the solvent-producing
 RT bacterium Clostridium acetobutylicum."
 RL J. Bacteriol. 183:4823-4838(2001).
 RN [2]

SEQUENCE OF 1-108 FROM N.A.
 RP STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RC MEDLINE=93273706; PubMed=8501044;
 RX Sauer U., Duerre P.;
 RA "Sequence and molecular characterization of a DNA region encoding a
 RT small heat shock protein of Clostridium acetobutylicum."
 RL J. Bacteriol. 175:3394-3400(1993).
 CC -!- SIMILARITY: BELONGS TO THE FTSK/SPOIIIE FAMILY.
 CC -!- CAUTION: Ref.2 sequence differs from that shown due to frameshifts
 CC in positions 76 and 106.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; A8007866; AAK81629.1; -
 DR EMBL; X65276; CAA46379.1; ALT FRAME.
 DR InterPro; IPR002543; FtsK_SpoIIIE.
 DR Pfam; PF01580; FtsK_SpoIIIE; 2.
 KW Hypothetical protein; ATP-binding; Complete proteome.
 FT NP BIND 675 682 ATP [POTENTIAL].
 SQ SEQUENCE 1498 AA; 168968 MW; FF42037A335A9649 CRC64;

Query Match 69.2%; Score 36; DB 1; Length 1498;
 Best Local Similarity 60.0%; Pred. No. 23;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 BEVVPXGMSY 10
 : :|||
 Db 1276 EQKIPNGMSY 1285

RESULT 5
 ZEPI_HUMAN STANDARD; PRT; 2717 AA.
 ID ZEPI_HUMAN
 AC P15822;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Zinc finger protein 40 (Human immunodeficiency virus type I enhancer-
 DE binding protein 1) (HIV-EPI) (Major histocompatibility complex binding
 DE protein 1) (MBP-1) (Positive regulatory domain II binding factor 1)
 DE (PRDII-BF1).
 DE HIVEPI OR ZNF40.
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90169514; PubMed=2106471;
 RA Fan C.M., Maniatis T.;
 RT "A DNA-binding protein containing two widely separated zinc finger
 RT motifs that recognize the same DNA sequence."
 RL Genes Dev. 4:29-42(1990).
 RN [2]

STRUCTURE BY NMR OF 2113-2142.
 RX MEDLINE=91064333; PubMed=2248949;
 RA Omichinski J.G., Clore G.M., Appella E., Sakaguchi K.,
 RA Gronenborn A.M.;
 RT "High-resolution three-dimensional structure of a single zinc finger
 RT from a human enhancer binding protein in solution."
 RL Biochemistry 29:9324-9334(1990).
 RN [3]

STRUCTURE BY NMR OF 2087-2142.
 RX MEDLINE=92232684; PubMed=1567844;
 RA Omichinski J.G., Clore G.M., Robien M., Sakaguchi K., Appella E.,
 RA Gronenborn A.M.;
 RT "High-resolution solution structure of the double Cys2His2 zinc
 RT finger from the human enhancer binding protein MBP-1."
 RL Biochemistry 31:3907-3917(1992).

CC -!- FUNCTION: THIS PROTEIN SPECIFICALLY BINDS TO THE DNA SEQUENCE
 CC 5'-GGGACTTTC-3' WHICH IS FOUND IN THE ENHANCER ELEMENTS OF
 CC NUMEROUS VIRAL PROMOTERS SUCH AS THOSE OF SV40, CMV, OR HIV1.
 CC IN ADDITION, RELATED SEQUENCES ARE FOUND IN THE ENHANCER ELEMENTS
 CC OF A NUMBER OF CELLULAR PROMOTERS, INCLUDING THOSE OF THE CLASS I
 CC MHC, INTERLEUKIN-2 RECEPTOR, AND INTERFERON-BETA GENES. IT MAY ACT
 CC IN T-CELL ACTIVATION.

CC -!- SUBCELLULAR LOCATION: Nuclear.

CC -!- INDUCTION: BY MITOGEN AND PHORBOL ESTER.

CC -!- DOMAIN: CONTAINS TWO SETS OF 2 ZINC-FINGERS, WHICH ARE WIDELY
 CC SEPARATED AND RECOGNIZE THE SAME DNA SEQUENCE. THERE IS A FIFTH
 CC ZINC-FINGER IN-BETWEEN.

CC -!- SIMILARITY: STRONG, TO HIVEP2.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X51435; CAA35798.1; -
 DR PIR; A34203; A34203.
 DR PDB; 3ZNF; 15-JAN-92.
 DR PDB; 4ZNF; 15-JAN-92.
 DR PDB; 1BBO; 31-OCT-93.
 DR TRANSFAC; T00497; -

```

DR Genew; HGNC:4920; HIVP1.
DR MIM; 194540; -.
DR InterPro; IPR000822; Znf_C2H2; 5.
DR Pfam; PF00096; zf-C2H2; 5.
DR PRINTS; PRO0048; ZINC_FINGER.
DR SMART; SM00355; Znf_C2H2; 4.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 4.
KW Transcription regulation; Zinc-finger; Metal-binding; DNA-binding;
KW Nuclear protein; Repeat; 3D-structure.
FT DOMAIN 406 456
FT ZN_FING 406 428 ZINC FINGERS.
FT ZN_FING 434 456 C2H2-TYPE.
FT ZN_FING 434 456 C2H2-TYPE.
FT DOMAIN 803 806 POLY-SER.
FT ZN_FING 958 981 C2HC-TYPE (POTENTIAL).
FT DOMAIN 2087 2139 ZINC FINGERS.
FT ZN_FING 2087 2109 ZINC FINGERS.
FT ZN_FING 2115 2139 C2H2-TYPE.
FT STRAND 2088 2088 C2H2-TYPE.
FT TURN 2090 2092
FT STRAND 2095 2095
FT STRAND 2095 2095
FT HELIX 2099 2108
FT TURN 2109 2109
FT STRAND 2115 2116
FT STRAND 2123 2124
FT HELIX 2127 2135
SQ SEQUENCE 2717 AA; 297217 MW; D45D3CA951FEA561 CRC64;

Query Match 67.3%; Score 35; DB 1; Length 2717;
Best Local Similarity 66.7%; Pred. No. 70;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11
DB 2405 VVPAGLTYS 2413

RESULT 6
CY14_NEUCR
ID CY14_NEUCR STANDARD; PRT; 788 AA.
AC F23622;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sulfate permease II.
GN CYS-14.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91129256; PubMed=1825178;
RA Ketter J.S., Jarai G., Fu Y.-H., Marzluf G.A.;
RT "Nucleotide sequence, messenger RNA stability, and DNA recognition
RT elements of cys-14, the structural gene for sulfate permease II in
RT Neurospora crassa";
RL Biochemistry 30:1780-1787(1991).
RN [2]
RP PROBABLE REVISIONS.
RX MEDLINE=9418926; PubMed=8140616;
RA Sandal N.N., Marcker K.A.;
RT "Similarities between a soybean nodulin, Neurospora crassa sulphate
RT permease II and a putative human tumour suppressor.";
RL Trends Biochem. Sci. 19:19-19(1994).
CC -!- FUNCTION: UPTAKE OF SULFATE INTO THE CELL.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- INDUCTION: Highly expressed, but only in cells subject to sulfur
CC limitation, and it is turned on by the positive-acting Cys-3
CC sulfur regulatory protein.
CC -!- MISCELLANEOUS: SULFATE PERMEASE II IS MAINLY FOUND IN MYCELIA.
CC -!- SIMILARITY: BELONGS TO THE SLC26A FAMILY OF TRANSPORTERS.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M59167; AAA33615.1; ALT_SEQ.
DR PIR; A37956; A37956. Sulfate_transp.
DR InterPro; IPR001902; Sulfate_transp; 1.
DR Pfam; PF00916; Sulfate_transp; 1.
DR TIGRfams; TIGR00815; sulp; 1.
DR PROSITE; PS01130; SLC26A; 1.
KW Transport; Transmembrane; Glycoprotein.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 103 123 POTENTIAL.
FT TRANSMEM 128 148 POTENTIAL.
FT TRANSMEM 171 191 POTENTIAL.
FT TRANSMEM 193 213 POTENTIAL.
FT TRANSMEM 271 291 POTENTIAL.
FT TRANSMEM 326 346 POTENTIAL.
FT TRANSMEM 363 383 POTENTIAL.
FT TRANSMEM 451 471 POTENTIAL.
FT TRANSMEM 474 494 POTENTIAL.
FT CARBOHYD 23 23 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 578 578 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 788 AA; 87864 MW; 4FC604B60798CE77 CRC64;

Query Match 65.4%; Score 34; DB 1; Length 788;
Best Local Similarity 66.7%; Pred. No. 32;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11
DB 90 VVPQGMAYA 98

RESULT 7
AL0C_HUMAN
ID AL0C_HUMAN STANDARD; PRT; 1499 AA.
AC O60312; Q96914;
DT 30-MAY-2000 (Rel. 39, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Potential phospholipid-transporting ATPase VC (EC 3.6.3.1) (ATPVC)
DE (Aminophospholipid translocase VC).
GN ATP10C OR ATPVC OR KIAA0566.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21225279; PubMed=11326269;
RA Meguro M., Kashiwagi A., Mitsuya K., Nakao M., Kondo I., Saitoh S.,
RA Oshimura M.;
RT "A novel maternally expressed gene, ATP10C, encodes a putative
RT aminophospholipid translocase associated with Angelman syndrome.";
RL Nat. Genet. 28:19-20(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21313119; PubMed=11353404;
RA Herzig L.B.K., Kim S.-J., Cook E.H. Jr., Ledbetter D.H.;
RT "The human aminophospholipid-transporting ATPase gene ATP10C maps
RT adjacent to UBE3A and exhibits similar imprinted expression.";
RL Am. J. Hum. Genet. 68:1501-1505(2001).
RN [3]
RP SEQUENCE OF 337-1499 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96290545; PubMed=9628581;
RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
RA Nomura N., Ohara O.;

```

"Prediction of the coding sequences of unidentified human genes. IX. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro.";

CC -!- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).

CC -!- TISSUE SPECIFICITY: Wide expression, with highest levels in kidney, followed by lung, brain, prostate, testis, ovary, and small intestine.

CC -!- DISEASE: Defects in ATP10C may be a cause of Angelman syndrome (AS), also known as 'happy puppet syndrome'.

CC -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY (E1-E2 ATPASES). SUBFAMILY IV.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AB051358; BAB47392.1; -.

DR EMBL; AY029504; AAK33100.1; -.

DR EMBL; AY029487; AAK33100.1; JOINED.

DR EMBL; AY029488; AAK33100.1; JOINED.

DR EMBL; AY029489; AAK33100.1; JOINED.

DR EMBL; AY029490; AAK33100.1; JOINED.

DR EMBL; AY029491; AAK33100.1; JOINED.

DR EMBL; AY029492; AAK33100.1; JOINED.

DR EMBL; AY029493; AAK33100.1; JOINED.

DR EMBL; AY029494; AAK33100.1; JOINED.

DR EMBL; AY029495; AAK33100.1; JOINED.

DR EMBL; AY029496; AAK33100.1; JOINED.

DR EMBL; AY029497; AAK33100.1; JOINED.

DR EMBL; AY029498; AAK33100.1; JOINED.

DR EMBL; AY029499; AAK33100.1; JOINED.

DR EMBL; AY029500; AAK33100.1; JOINED.

DR EMBL; AY029501; AAK33100.1; JOINED.

DR EMBL; AY029502; AAK33100.1; JOINED.

DR EMBL; AY029503; AAK33100.1; JOINED.

DR EMBL; AB011138; BAA25492.1; -.

DR GENE; HGNC:13547; ATP10C.

DR MIM; 605855; -.

DR MIM; 105830; -.

DR InterPro; IPR001757; ATPase_E1-E2.

DR InterPro; IPR001454; Hlgnaase/hydrase.

DR Pfam; PF00702; Hydrolase; 1.

DR PRINTS; PR001119; CATAPASE.

DR PROSITE; PS00154; ATPase E1 E2; 1.

KW Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding; Multigene family.

KW DOMAIN 1 86 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 87 106 POTENTIAL.

FT DOMAIN 107 110 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 111 128 POTENTIAL.

FT DOMAIN 129 309 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 310 332 POTENTIAL.

FT DOMAIN 337 362 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 363 384 POTENTIAL.

FT DOMAIN 385 1087 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 1088 1108 POTENTIAL.

FT DOMAIN 1109 1119 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 1120 1140 POTENTIAL.

FT DOMAIN 1141 1170 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 1171 1192 EXTRACELLULAR (POTENTIAL).

FT DOMAIN 1193 1199 POTENTIAL.

FT TRANSMEM 1200 1222 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 1223 1228 POTENTIAL.

FT TRANSMEM 1229 1249 EXTRACELLULAR (POTENTIAL).

FT DOMAIN 1250 1267 POTENTIAL.

FT TRANSMEM 1268 1292 EXTRACELLULAR (POTENTIAL).

FT DOMAIN 1293 1499 CYTOPLASMIC (POTENTIAL).

FT MOD_RES 427 427 PHOSPHORYLATION (BY SIMILARITY).

FT METAL 1031 1031 MAGNESIUM (BY SIMILARITY).

FT METAL 1035 1035 MAGNESIUM (BY SIMILARITY).

FT DOMAIN 467 470 POLY-GLU.

FT CONFLICT 388 388 Q -> R (IN REF. 3).

SQ SEQUENCE 1499 AA; 167687 MW; D4996A4D0635A68D CRC64;

Query Match 65.4%; Score 34; DB 1; Length 1499;

Best Local Similarity 72.7%; Pred. No. 62;

Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11

Db 469 EEVVPXGMSYS 11

RESULT 8

RL20_MYCPU STANDARD; PRT; 116 AA.

AC Q98QV0;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE 50S ribosomal protein L20.

GN RPLT OR MYPU 2610.

OS Mycoplasma pulmonis.

OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

OX NCBI_TaxID=2107;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=UAB CTIP;

RX MEDLINE=21267165; PubMed=11353084;

RA Chambaud I., Heilig R., Ferris S., Barbe V., Samson D., Galisson F., Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P.C., Blanchard A.

RA "The complete genome sequence of the murine respiratory pathogen RT Mycoplasma pulmonis.";

RL Nucleic Acids Res. 29:2145-2153(2001).

CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA AND IS NECESSARY TO THE IN VITRO ASSEMBLY PROCESS OF THE 50S RIBOSOMAL SUBUNIT; IT IS NOT INVOLVED IN THE PROTEIN SYNTHESIZING FUNCTIONS OF THAT SUBUNIT (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE L20P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AL445563; CAC13434.1; -.

DR Mypulist; MYPU_2610; -.

DR InterPro; IPR001081; Ribosomal_L20.

DR Pfam; PF00453; Ribosomal_L20; 1.

DR PRINTS; PR00062; RIBOSOMALL20.

DR ProDom; PD002389; Ribosomal_L20; 1.

DR TIGRfams; TIGR01032; rplT_bact; 1.

DR PROSITE; PS00937; RIBOSOMAL_L20; 1.

KW Ribosomal protein; rRNA-binding; Complete proteome.

SQ SEQUENCE 116 AA; 13565 MW; C59C748901B18F14 CRC64;

Query Match 63.5%; Score 33; DB 1; Length 116;

Best Local Similarity 77.8%; Pred. No. 7;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVFXGMSYS 11

Db 68 VVFXGMSYS 76

RESULT 9

```

Y990 CAMJE
ID _Y990 CAMJE STANDARD; PRT; 253 AA.
AC P45489; O9PNV0;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein CJ0990C.
GN CJ0990C.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 11168;
RA MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrett B.G.;
RA "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences.";
RL Nature 403:665-668(2000).
RN [2]
RP SEQUENCE OF 160-253 FROM N.A.
RC STRAIN=ATCC 43431 / TGH 9011;
RX MEDLINE=95247673; PubMed=7730270;
RA Hani E.K., Chan V.L.;
RT "Expression and characterization of Campylobacter jejuni
RT benzoylglycine amidohydrolase (Hippuricase) gene in Escherichia
RL coli.";
RN J. Bacteriol. 177:2396-2402(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL139076; CAB73246.1; -.
DR EMBL; Z36940; CA85392.1; -.
RW Hypothetical protein; Complete proteome.
SQ SEQUENCE 253 AA; 29783 MW; F96D3FF3265F8A6A CRC64;

Query Match 63.5%; Score 33; DB 1; Length 253;
Best Local Similarity 55.6%; Pred. No. 16;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 EVVPXGMSY 10
Db 185 DIPFGMSY 193
: : |||||
: : |||||

RESULT 10
AROA VIBCH STANDARD; PRT; 426 AA.
AC O9KRE0;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE enolpyruvylshikimate-3-phosphate synthase) (BPS synthase) (BPSPS).
GN AROA OR VC1732.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;

```

```

RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
CC -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC sixth step.
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE004251; AAF94882.1; -.
DR TIGR; VC1732; -.
DR InterPro; IPR001986; EPSP_synthase.
DR Pfam; PF00275; EPSP_synthase; 1.
DR ProDom; PD001867; EPSP_synthase; 1.
DR PROSITE; PS00104; EPSP_SYNTHASE 1; 1.
DR PROSITE; PS00885; EPSP_SYNTHASE 2; 1.
KW Aromatic amino acid biosynthesis; Transferase; Complete proteome.
SQ SEQUENCE 426 AA; 46101 MW; 38852D6483BFE1C3 CRC64;

Query Match 63.5%; Score 33; DB 1; Length 426;
Best Local Similarity 60.0%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EVVPXGMSY 10
Db 223 EFVIPAGQSY 232
: : |||||
: : |||||

RESULT 11
CC37 SCHPO STANDARD; PRT; 466 AA.
AC O94740;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hsp90 co-chaperone Cdc37 (Hsp90 chaperone protein kinase-targeting
DE subunit) (Cell division control protein 37).
GN CDC37 OR SPAC9B6.10.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC Westwood P.K., Preston N.C., Fantes P.A.;
RA "Schizosaccharomycetes pombe cdc37 gene.";
RT Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,

```


RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Welljens I., Vanstaerck E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabi C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leinach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
RT "The genome sequence of Schizosaccharomyces pombe";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: With Hsp90 it forms a complex that binds to several
CC kinases, resulting in stabilization and promotion of their
CC activity (by similarity).
CC -!- SUBUNIT: Forms a complex with Hsp90. Interacts with a number of
CC kinases (by similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE CDC37 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AJ132377; CAB38758.1; -;
CC DR EMBL; AJ132376; CAB38757.1; -;
CC DR EMBL; AL049769; CAB42371.2; -;
CC KW Chaperone; Cell division; Cell cycle.
CC SQ SEQUENCE 466 AA; 52554 MW; 64723B8B34CABB3C5 CRC64;

Query Match 63.5%; Score 33; DB 1; Length 466;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVVFXGMSY 10
Db : : |||||
98 DSAIPGMSY 107

RESULT 12
GSR2 HUMAN STANDARD; PRT; 478 AA.
AC Q9NZM5; Q9NPPI; Q9UP12; Q9BTC6; Q9HAX6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glioma tumor suppressor candidate region gene 2 protein (p60).
GN GLTSCR2.
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX MEDLINE=20175430; PubMed=10708517;
RA Smith J.S., Tachibana I., Pohl U., Lee H.K., Thanarajasingam U.,
RA Portier B.P., Ueki K., Billings S., Ramaswamy S., Mohrenweiser H.W.,
RA Scheithauer B.W., Louis D.N., Jenkins R.B.;
RT "A transcript map of the chromosome 19q-Arm glioma tumor suppressor

region."; Genomics 64:44-50(2000).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE OF 9-478 FROM N.A., AND SUBCELLULAR LOCATION.
RX MEDLINE=99214318; PubMed=10196275;
RA Bruni R., Fineschi B., Ogle W.O., Roizman B.;
RT "A novel cellular protein, p60, interacting with both herpes simplex
RT virus 1 regulatory proteins ICP22 and ICP0 is modified in a
RT cell-type-specific manner and is recruited to the nucleus after
RT infection."; J. Virol. 73:3810-3817(1999).
[4]
RP SEQUENCE OF 12-478 FROM N.A.
RA Andreu N., Estivill X., Escarceller M., Sumoy L.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
[5]
RP SEQUENCE OF 218-477 FROM N.A.
RC TISSUE=Testis;
RA Poustka A., Klein M., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: Interacts with HSV-1 early proteins ICP22 and ICP0.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Expressed at high levels in heart and
CC pancreas, moderate levels in placenta, liver, skeletal muscle, and
CC kidney, and low levels in brain and lung.
CC -!- SIMILARITY: BELONGS TO THE GLTSCR2 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AF182076; AAF62873.1; -;
CC DR EMBL; BC004229; AAH04229.1; -;
CC DR EMBL; BC006311; AAH06311.1; -;
CC DR EMBL; BC010095; AAH10095.1; -;
CC DR EMBL; AF296124; AAG30413.1; -;
CC DR EMBL; AL359335; CAB94786.1; -;
CC DR EMBL; AL359336; CAB94787.1; -;
CC DR EMBL; AL122063; CAB59242.1; -;
CC DR SWISS-2DPAGE; Q9NZM5; HUMAN.
CC Genew; HGNC:4333; GLTSCR2.
CC MIM; 605691; -;
CC KW Nuclear protein; Polymorphism.
CC VARIANT 389 389 R -> Q.
FT CONFLICT 4 6 /FTID=VAR 011486.
FT CONFLICT 9 9 GGS -> HEG (IN REF. 2; AAH04229).
FT CONFLICT 146 191 TVERP -> SGRSSYGRSPRASSPGGQSPVAQPCN
FT CONFLICT 198 215 KGNPAPGHRIAA (IN REF. 3).
FT CONFLICT 235 235 SDNPLDPLVGGQDEFFLE -> LNNPDKPVVWPGCLFPG
FT CONFLICT 417 417 A -> S (IN REF. 2; AAH04229).
FT CONFLICT 433 477 PEGNILDREFKSFQRNMIEPRERAKFKKVKYKLVKEKAF
FT CONFLICT 434 478 REIQ -> VLTVSCRGAPCPVMTPLPVPFPGYGHGCGP
FT CONFLICT 478 478 WAGVGPMPRG (IN REF. 5).
FT CONFLICT 478 478 EGNILDREFKSFQRNMIEPRERAKFKKVKYKLVKEKAF
FT CONFLICT 478 478 EIQIL -> RQHSFETGSAFRGGI (IN REF. 3).
SQ SEQUENCE 478 AA; 54417 MW; 7F18923E348CB52B CRC64;

Query Match 63.5%; Score 33; DB 1; Length 478;
Best Local Similarity 60.0%; Pred. No. 30;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 EVVFXGMSYS 11
 ||| |||
 Db 239 EVAPAGASYN 248

RESULT 13

BCN5 CLOPE
 ID BCN5 CLOPE STANDARD; PRT; 890 AA.
 AC P08656;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE Bacteriocin BCN5.
 GN BCN.
 OS Clostridium perfringens.
 OG Plasmid pIP404.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1502;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CPN50;
 RX MEDLINE=88336297; PubMed=2901769;
 RA Garnier T., Cole S.T.;
 RT "Complete nucleotide sequence and genetic organization of the
 RT bacteriocinogenic plasmid, pIP404, from Clostridium perfringens.";
 RL Plasmid 19:134-150(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CPN50;
 RX MEDLINE=87057020; PubMed=2877971;
 RA Garnier T., Cole S.T.;
 RT "Characterization of a bacteriocinogenic plasmid from Clostridium
 RT perfringens and molecular genetic analysis of the
 RT bacteriocin-encoding gene.";
 RL J. Bacteriol. 168:1189-1196(1986).
 RN [3]
 RP SEQUENCE OF 1-14 FROM N.A.
 RC STRAIN=CPN50;
 RX MEDLINE=89039249; PubMed=2460717;
 RA Garnier T., Cole S.T.;
 RT "Studies of UV-inducible promoters from Clostridium perfringens in
 RT vivo and in vitro.";
 RL Mol. Microbiol. 2:607-614(1988).
 CC -1- FUNCTION: MAY FUNCTION AS AN IONOPHORE.
 CC -1- INDUCTION: BY UV IRRADIATION.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).

EMBL; M14481; AAA98248.1; -.
 PIR; M32882; AAA98249.1; -.
 DR A30481; A30481.
 DR InterPro; IPR003646; SH3_bac.
 DR SMART; SM00287; SH3b; 3.
 KW Antibiotic; Bacteriocin; Plasmid.
 FT DOMAIN 815 869 HYDROPHOBIC.
 SQ SEQUENCE 890 AA; 96699 MW; F4E5E8971C31C6C6 CRC64;

Query Match 63.5%; Score 33; DB 1; Length 890;
 Best Local Similarity 66.7%; Pred. No. 58;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 EVVFXGMSYS 10
 ||| |||
 Db 170 EVVPGGFTY 178

RESULT 14
 LTBS HUMAN
 ID LTBS HUMAN STANDARD; PRT; 1394 AA.
 AC P22064;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Latent transforming growth factor beta binding protein 1S precursor
 DE (Transforming growth factor beta-1 binding protein 1) (TGF-beta1-BP-
 DE 1).
 GN LTBP1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Fibroblast, and Platelet;
 RX MEDLINE=90275601; PubMed=2350783;
 RA Kanazaki T., Olofsson A., Moren A., Wernstedt C., Hellman U.,
 RA Miyazono K., Claesson-Welsh L., Heldin C.-H.;
 RT "TGF-beta 1 binding protein: a component of the large latent complex
 RT of TGF-beta 1 with multiple repeat sequences.";
 RL Cell 61:1051-1061(1990).
 CC -1- SUBUNIT: THE LARGE LATENT COMPLEX OF TGF-BETA1 FROM PLATELETS IS
 CC COMPOSED OF THE TGF-BETA1 MOLECULE NONCOVALENTLY ASSOCIATED WITH
 CC A DISULFIDE-BONDED COMPLEX OF A DIMER OF THE N-TERMINAL PROPEPTIDE
 CC OF THE TGF-BETA1 PRECURSOR AND A THIRD COMPONENT DENOTED TGF-
 CC BETA1-BP. TGF-BETA1-BP DOES NOT BIND DIRECTLY TO ACTIVE TGF-BETA1.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a short form (shown here) and a
 CC long form (AC Q14766); are produced by alternative splicing.
 CC -1- PTM: CONTAINS HYDROXYLATED ASPARAGINE RESIDUES.
 CC -1- PTM: The N-terminus is blocked.
 CC -1- SIMILARITY: CONTAINS 16 EGF-LIKE DOMAINS.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).

EMBL; M34057; AAA61160.1; -.

PIR; A35626; A35626.

HSSP; P00750; ITPG.

GlycoSuiteDB; P22064; -.

Genew; HGNC:6714; LTBP1.

MM; 150390; -.

InterPro; IPR000152; Asx hydroxyl.

InterPro; IPR000561; EGF-like.

InterPro; IPR001881; EGF_Ca.

InterPro; IPR002212; Fibril-associ.

Pfam; PF00008; EGF; 15.

Pfam; PF00683; TB; 4.

SMART; SM00179; EGF_Ca; 13.

SMART; SM00001; EGF-like; 4.

PROSITE; PS00010; ASX_HYDROXYL; 13.

PROSITE; PS00022; EGF_1; 2.

PROSITE; PS01186; EGF_2; 11.

PROSITE; PS01187; EGF_Ca; 15.

Growth factor binding; Repeat; EGF-like domain; Hydroxylation; Signal;

Glycoprotein; Alternative splicing.

FT SIGNAL 1 20

FT CHAIN 21 1394

POTENTIAL

LATENT TRANSFORMING GROWTH FACTOR BETA

BINDING PROTEIN 1S.

EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).

REPEAT A.

EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).

EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).

EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).

EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).

DOMAIN 300 340

DOMAIN 348 412

DOMAIN 546 587

DOMAIN 588 629

DOMAIN 630 671

DOMAIN 671 710

	Query Match	63.5%;	Score 33;	DB 1;	Length 1394;
	Best Local Similarity	45.5%;	Pred. No. 93;		
	Matches	5; Conservative	3; Mismatches	3; Indels	0; Gaps
Dy	1 EVVVPXGMSYS 11	:	:	:	:
	: : :	:	:	:	:
Ddb	399 KEICPGMGYT 409	:	:	:	:
	STANDARD;	PRT;	1401 AA.		
	RPOC_VIBCH				
	Q9KV23;				
	15-JUN-2002 (Rel. 41, Created)				
	15-JUN-2002 (Rel. 41, Last sequence update)				
	15-JUN-2002 (Rel. 41, Last annotation update)				
	DNA-directed RNA polymerase beta' chain (EC 2.7.7.6) (Transcriptase				
	'beta' chain) (RNA polymerase beta' subunit).				
	RPOC OR VC0329.				
	Vibrio cholerae.				
	Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.				
	NCBI_TaxID=666;				
	[1]				
	SEQUENCE FROM N.A.				
	STRAIN=El Tor N16961 / Serotype O1;				
	MEDLINE=20406833; PubMed=10952301;				
	Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,				
	Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,				
	Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,				
	Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,				
	McDonald L., Uterback T., Fleischmann R.D., Nierman W.C., White O.,				
	Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,				
	Fraser C.M.;				
	"DNA sequence of both chromosomes of the cholera pathogen Vibrio				
	cholerae.";				
	Nature 406:477-483(2000).				
	-!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION				
	OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS				
	SUBSTRATES (BY SIMILARITY).				
	-!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +				
	{RNA} (N).				
	-!- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE				
	ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1				
	BETA' CHAIN (BY SIMILARITY).				
	-!- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.				

	This SWISS-PROT entry is copyright. It is produced through a collaboration				
	between the Swiss Institute of Bioinformatics and the EMBL outstation				
	the European Bioinformatics Institute. There are no restrictions on its				
	use by non-profit institutions as long as its content is in no way				
	modified and this statement is not removed. Usage by and for commercial				
	entities requires a license agreement (See http://www.isb-sib.ch/announce				
	or send an email to license@isb-sib.ch).				

	EMBL; AE004121; AAF93502.1; -				
	HSSP; Q9KWD6; 1HQW.				
	TIGR; VC0329; -				
	InterPro; IPR000722; RNA_pol_A.				
	InterPro; IPR002879; RNA_pol_A2.				
	Pfam; PF00623; RNA_pol_A; 1.				
	Pfam; PF01854; RNA_pol_A2; 2.				
	Transferrase; DNA-directed RNA polymerase; Transcription;				
	Complete proteome.				
	SEQUENCE 1401 AA; 155021 MW; DFDB0F2B5514504F CRC64;				

	Query Match	63.5%;	Score 33;	DB 1;	Length 1401;
	Best Local Similarity	50.0%;	Pred. No. 93;		
	Matches	5; Conservative	3; Mismatches	2; Indels	0; Gaps
Qy	2 EVVVPXGMSYS 11	:	:	:	:
	: : :	:	:	:	:
Ddb	581 QIVPKGLPYIS 590	:	:	:	:

Search completed: June 4, 2003, 13:11:44
Job time : 7.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:08:49 ; Search time 11 Seconds
(without alignments)
96.134 Million cell updates/sec

Title: AUDET-909-1

Perfect score: 52

Sequence: 1 eevvpqgmsys 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	73.1	3472	2 T31308	hypothetical 367K
2	37	71.2	840	2 T39116	probable sulfate p
3	37	71.2	877	2 T40413	sulfate permease -
4	36	69.2	102	2 A42452	VI protein - tobac
5	36	69.2	1498	2 B37355	DNA segregation AT
6	35	67.3	225	2 S57810	hypothetical prote
7	35	67.3	425	2 T24111	hypothetical prote
8	35	67.3	670	2 S22293	zinc finger protei
9	35	67.3	749	2 H82691	topoisomerase IV s
10	35	67.3	2717	2 A34203	DNA-binding protei
11	34	65.4	156	2 S54619	hypothetical prote
12	34	65.4	252	2 H69491	cell division inhi
13	34	65.4	544	2 C92900	probable ABC subst
14	33	63.5	94	2 I40758	hypothetical prote
15	33	63.5	116	2 E30544	50S ribosomal prot
16	33	63.5	165	2 D69493	hypothetical prote
17	33	63.5	253	2 C81374	hypothetical prote
18	33	63.5	259	2 T34536	hypothetical prote
19	33	63.5	284	2 S75817	hypothetical prote
20	33	63.5	298	2 T47670	beta-ketoacyl-ACP
21	33	63.5	368	2 F72281	hypothetical prote
22	33	63.5	426	2 D82163	3-phosphoshikimate
23	33	63.5	466	2 T33653	cdc37 protein - fl
24	33	63.5	653	2 D82352	iron(III) ABC tran
25	33	63.5	890	2 A30481	bacteriocin BCN5 -
26	33	63.5	1028	2 AF3286	ATP-dependent DNA
27	33	63.5	1152	2 D87046	conserved hypotet
28	33	63.5	1394	2 A35626	transforming growt
29	33	63.5	1401	2 G82336	DNA-directed RNA p

ALIGNMENTS

RESULT 1

T31308

hypothetical 367K protein - Cenarchaeum symbiosum

C:Species: Cenarchaeum symbiosum

C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 18-Feb-2000

C:Accession: T31308

R:Schleper, C.; DeLong, E.F.; Preston, C.M.; Feldman, R.A.; Wu, K.Y.; Swanson, R.V.

J. Bacteriol. 180, 5003-5009, 1998

A>Title: Genomic analysis reveals chromosomal variation in natural populations of the unc

A:Reference number: Z20994; MUID:98422450; PMID:9748430

A:Accession: T31308

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-3472 <SCH>

A:Cross-references: EMBL:AF083072; NID:G3599393; PID:G3599394; PIDN:AAC62699.1

C:Superfamily: Cenarchaeum symbiosum hypothetical 367K protein

Query Match 73.1%; Score 38; DB 2; Length 3472;

Best Local Similarity 54.5%; Pred. No. 59;

Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11

Db 2294 EDVIPRGISFS 2304

RESULT 2

T39116

probable sulfate permease - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T39116

R:Hunt, C.; Aves, S.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, November 1999

A:Reference number: Z21829

A:Accession: T39116

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-840 <HUN>

A:Cross-references: EMBL:AL132779; PIDN:CAB60015.1; GSPDB:GN00066; SPDB:SPAC869.05C

C:Genetics: Experimental source: strain 972h; cosmid c869

A:Gene: SPDB:SPAC869.05C

A:Map position: 1

Query Match 71.2%; Score 37; DB 2; Length 840;

Best Local Similarity 77.8%; Pred. No. 21;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11

Db 135 VVPQMSYA 143

```
RESULT 3
T40413
sulfate permease - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T40413
R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Jimenez Martinez, J.
submitted to the EMBL Data Library, August 1998
A:Reference number: 221926
A:Accession: T40413
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-877 <LYN>
A:Cross-references: EMBL:AL031261; PIDN:CAA20298.1; GSPDB:GN000667; SPDB:SPBC3H7.02
A:Experimental source: strain 972h-; cosmid c3H7
C:Genetics:
A:Gene: SPDB:SPBC3H7.02
A:Map position: 2

Query Match          71.2%; Score 37; DB 2; Length 877;
Best Local Similarity 77.8%; Pred. No. 22;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11
    ||| |||||
Db 148 VVPQGSYA 156

RESULT 4
A42452
V1 protein - tobacco yellow dwarf virus (strain Australia)
C:Species: tobacco yellow dwarf virus
C>Date: 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change 08-Oct-1999
C:Accession: A42452
R:Morris, B.A.M.; Richardson, K.A.; Haley, A.; Zhan, X.; Thomas, J.E.
Virology 187, 633-642, 1992
A:Title: The nucleotide sequence of the infectious cloned DNA component of tobacco yellow
dwarf virus
A:Reference number: A42452; MUID:92188538; PMID:1546458
A:Accession: A42452
A:Molecule type: DNA
A:Residues: 1-102 <VOR>
A:Cross-references: GB:M81103; NID:G335283; PIDN:AAA7947.1; PID:G335284

Query Match          69.2%; Score 36; DB 2; Length 102;
Best Local Similarity 60.0%; Pred. No. 3.5;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11
    ||||| |||
Db 7 QVVPSGINYS 16

RESULT 5
B97355
DNA segregation ATPase, FtsK/SpoIIIE family, YUKA B. subtilis ortholog [imported] - Clos
C:Species: Clostridium acetobutylicum
C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: B97355
R:Nolling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
stridium acetobutylicum
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: B97355
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1498 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK81629.1; PID:GI5026814; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC3709
```

```
Query Match          69.2%; Score 36; DB 2; Length 1498;
Best Local Similarity 60.0%; Pred. No. 63;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 BEVVPXGMSY 10
    ||| |||||
Db 1276 EOKIPMGMSY 1285

RESULT 6
S57810
hypothetical protein precursor (clone TPp11) - tomato
C:Species: Lycopersicon esculentum (tomato)
C>Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jul-2000
C:Accession: S57810
R:Milligan, S.B.; Gasser, C.S.
Plant Mol. Biol. 28, 691-711, 1995
A:Title: Nature and regulation of pistil-expressed genes in tomato.
A:Reference number: S57808; MUID:95375233; PMID:7647301
A:Accession: S57810
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-225 <MIL>
A:Cross-references: EMBL:U20592; NID:G924625; PIDN:AAA80497.1; PID:G924626
C:Superfamily: plant Kunitz-type proteinase inhibitor

Query Match          67.3%; Score 35; DB 2; Length 225;
Best Local Similarity 54.5%; Pred. No. 13;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 BEVVPXGMSYS 11
    ||||| |||
Db 32 DEVVPNGKTYA 42

RESULT 7
T24111
hypothetical protein R10D12.10 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T24111
R:Percy, C.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19842
A:Accession: T24111
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-425 <WIL>
A:Cross-references: EMBL:Z81109; PIDN:CAB03241.1; GSPDB:GN000023; CESP:R10D12.10
A:Experimental source: clone R10D12
C:Genetics:
A:Gene: CESP:R10D12.10
A:Map position: 5
A:Introns: 23/3; 56/3; 113/3; 257/2

Query Match          67.3%; Score 35; DB 2; Length 425;
Best Local Similarity 50.0%; Pred. No. 26;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 BEVVPXGMSY 10
    ||| |||||
Db 335 EQIVPGGLQY 344

RESULT 8
S22293
zinc finger protein AT-BP2 - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 20-Sep-1999
C:Accession: S22293; I78656
R:Michelmore, C.; Traboni, C.; Cortese, R.
Nucleic Acids Res. 19, 141-147, 1991
```

A:Title: Isolation of two cDNAs encoding zinc finger proteins which bind to the alpha 1-
A:Reference number: I58280; MUID:91187610; PMID:1301405
A:Accession: S22293
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-670 <MT>
A:Cross-references: EMBL:X54250; NID:957519; PIDN:CAA38151.1; PID:g57520
A>Note: the authors did not translate the codon for residue 1
C:Superfamily: HIV-EP2 enhancer-binding protein
C:Keywords: DNA binding; transcription regulation; zinc finger

Query Match 67.3%; Score 35; DB 2; Length 670;
Best Local Similarity 66.7%; Pred. No. 43;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11
||| ||:|
Db 376 VVPAGLTYS 384

RESULT 9
H2691
topoisomerase IV subunit XF1353 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C:Accession: H2691
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A>Note: for a complete list of authors see reference number A59328 below
A:Accession: H2691
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-749 <SIM>
A:Cross-references: GB:AB003967; GB:AB003849; NID:99106347; PIDN:AAF84162.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, P.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre, H
as-Neco, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF1353
C:Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomerase
Query Match 67.3%; Score 35; DB 2; Length 749;
Best Local Similarity 77.8%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10
||| |||
Db 526 EVDPSGMSY 534

RESULT 10
A34203
DNA-binding protein PRDII-BF1 - human
N:Alternate names: major histocompatibility complex enhancer-binding protein 1
C:Species: Homo sapiens (man)
C:Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 20-Sep-1999
C:Accession: A34203; A34779
R:Fan, C.M.; Maniatis, T.
Genes Dev. 4, 29-42, 1990

A:Title: A DNA-binding protein containing two widely separated zinc finger motifs that re
A:Reference number: A34203; MUID:90169514; PMID:2106471
A:Accession: A34203
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2717 <FAN>
A:Cross-references: EMBL:X51435; NID:938017; PIDN:CAA35798.1; PID:g38018
R:Baldwin Jr., A.S.; LeClair, K.P.; Singh, H.; Sharp, P.A.
Mol. Cell. Biol. 10, 1406-1414, 1990
A:Title: A large protein containing zinc finger domains binds to related sequence element
A:Reference number: A34779; MUID:90205817; PMID:2108316
A:Accession: A34779
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 801-1072, 'N', 1074-1168, 'K', 1170-1225, 'V', 1227-1434, 'N', 1436-1607, 'I', 1609-161
A:Cross-references: GB:M32019
C:Superfamily: HIV-EP2 enhancer-binding protein
C:Keywords: DNA binding; transcription regulation; zinc finger

Query Match 67.3%; Score 35; DB 2; Length 2717;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11
||| ||:|
Db 2405 VVPAGLTYS 2413

RESULT 11
S54619
hypothetical protein YOR013w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein O2612; hypothetical protein YOL303.3
C:Species: Saccharomyces cerevisiae
C:Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 19-Apr-2002
C:Accession: S54619; S66879
R:de Haan, M.; Maarse, A.C.; Grivell, L.A.
submitted to the EMBL Data Library, May 1995
A:Reference number: S54617
A:Accession: S54619
A:Molecule type: DNA
A:Residues: 1-156 <DEH>
A:Cross-references: EMBL:X87331; NID:g1041652; PIDN:CAA60762.1; PID:g829123
R:de Haan, M.; Grivell, L.A.; Maarse, A.C.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S66877
A:Accession: S66879
A:Molecule type: DNA
A:Residues: 1-156 <DEW>
A:Cross-references: EMBL:Z74920; NID:g1420109; PIDN:CAA99201.1; PID:g1420111; MIPS:YOR01
A:Experimental source: strain S288C
C:Genetics:
A:Cross-references: SGD:S0005539
A:Map position: 15R
C:Superfamily: hypothetical protein YOR013w
Query Match 65.4%; Score 34; DB 2; Length 156;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10
||| |||
Db 50 EVMPLGMDY 58

RESULT 12
H69491
cell division inhibitor (minD-2) homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 26-Aug-1999
C:Accession: H69491
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson,
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon
A:Reference number: A69250; MUID:98049343; PMID:9389475

A:Accession: H69491
A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-252 <KLE>
A:Cross-references: GB:AE000970; GB:AE000782; MID:g2689293; PIDN:AA89318.1; PID:g264860
C:Superfamily: cell division inhibitor mind

Query Match 65.4%; Score 34; DB 2; Length 252;

Best Local Similarity 75.0%; Pred. No. 24;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EVVPXGMS 9

Db 81 EVIPAGMS 88

RESULT 13

C82900

Probable ABC substrate-binding protein, iron U0359 [imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000

C:Accession: C82900

R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to Genbank, February 2000

A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min

A:Reference number: A82870

A:Accession: C82900

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-544 <GLA>

A:Cross-references: GB:AE002133; GB:AF222894; MID:g6899339; PIDN:AAF30768.1; GSPDB:GN001

A:Experimental source: serovar 3; biovar 1

C:Genetics:

A:Gene: ABCsbp-5; U0359

A:Genetic code: SGC3

Query Match

Best Local Similarity 65.4%; Score 34; DB 2; Length 544;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10

Db 135 EEVVPXGMSY 144

RESULT 14

I40758

hypothetical protein 1 - Campylobacter jejuni (fragment)

C:Species: Campylobacter jejuni

C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 08-Oct-1999

C:Accession: I40758; S47317

R:Hani, E.K.; Chan, V.L.

J. Bacteriol. 177, 2396-2402, 1995

A:Title: Expression and characterization of Campylobacter jejuni benzoylglycine amidohyd

A:Reference number: I40758; MUID:95247673; PMID:7730270

A:Accession: I40758

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-94 <RES>

A:Cross-references: EMBL:236940; MID:g535805; PIDN:CAA85392.1; PID:g535806

Query Match

Best Local Similarity 63.5%; Score 33; DB 2; Length 94;

Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 EVVPXGMSY 10

Db 26 DFFPGMSY 34

RESULT 15

E90544

50S ribosomal protein L20 [imported] - Mycoplasma pulmonis (strain UAB CTIP)
C:Species: Mycoplasma pulmonis

C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001

C:Accession: E90544

R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001

A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm

A:Reference number: A99512; MUID:21267165; PMID:11353084

A:Accession: E90544

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-116 <KUR>

A:Cross-references: GB:AL445566; PID:g14089674; PIDN:CAC13434.1; GSPDB:GN00153

A:Experimental source: strain UAB CTIP

C:Genetics:

A:Gene: MYPU 2610

A:Genetic code: SGC3

C:Superfamily: Escherichia coli ribosomal protein L20

Query Match

Best Local Similarity 63.5%; Score 33; DB 2; Length 116;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 VVPXGMSYS 11

Db 68 VRPLGMSYS 76

Search completed: June 4, 2003, 13:14:13

Job time : 12 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:11:19 ; Search time 183 seconds
(without alignments)
38.754 Million cell updates/sec

Title: AUDET-909-1

Perfect score: 52

Sequence: 1 eevpxgmssys 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main:*

- 1: /cgn2_6/ptodata/1/paa/US06 COMB.pcp.*
- 2: /cgn2_6/ptodata/1/paa/US06 COMB.pcp.*
- 3: /cgn2_6/ptodata/1/paa/US07 COMB.pcp.*
- 4: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 5: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 6: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 7: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 8: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 9: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 10: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 11: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 12: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 13: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 14: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*
- 15: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*
- 16: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*
- 17: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*
- 18: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*
- 19: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*
- 20: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*
- 21: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*
- 22: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*
- 23: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*
- 24: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*
- 25: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*
- 26: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*
- 27: /cgn2_6/ptodata/1/paa/US09 COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	76.9	1022	1	PCT-US01-08631-33980
2	40	76.9	1022	1	PCT-US01-08631-36185
3	40	76.9	1022	1	PCT-US01-08631-38532
4	38	73.1	3472	18	US-09-408-020-4
5	38	73.1	3472	24	US-10-027-801-4
6	38	73.1	3472	24	US-10-027-806-4

7	73.1	3472	24	US-10-029-120-4	Sequence 4, Appli
8	73.1	3472	24	US-10-034-623-4	Sequence 4, Appli
9	69.2	129	27	US-60-196-710-6207	Sequence 6207, Ap
10	69.2	140	27	US-60-177-646-2588	Sequence 2588, Ap
11	69.2	153	24	US-10-029-386-32486	Sequence 32486, A
12	69.2	382	15	US-09-134-000-3738	Sequence 3738, Ap
13	69.2	401	27	US-60-389-987-2598	Sequence 2598, Ap
14	69.2	401	27	US-60-412-418-2598	Sequence 2598, Ap
15	67.3	160	26	US-10-219-999-59270	Sequence 58270, A
16	67.3	160	27	US-60-312-544-9551	Sequence 9551, Ap
17	67.3	225	21	US-09-791-537-14785	Sequence 14785, A
18	67.3	336	24	US-10-029-386-32076	Sequence 32076, A
19	67.3	361	21	US-09-760-446A-1744	Sequence 1744, Ap
20	67.3	361	26	US-10-206-664-1744	Sequence 1744, Ap
21	67.3	425	21	US-09-791-537-95581	Sequence 95581, A
22	67.3	670	21	US-09-791-537-144347	Sequence 144347,
23	65.4	206	1	PCT-US02-30474-2302	Sequence 2302, Ap
24	65.4	206	27	US-60-324-631-2811	Sequence 2811, Ap
25	65.4	252	21	US-09-791-537-132485	Sequence 132485,
26	65.4	254	25	US-10-104-047-3921	Sequence 3921, Ap
27	65.4	387	15	US-09-181-375-2	Sequence 2, Appli
28	65.4	424	27	US-60-230-445-1331	Sequence 1331, Ap
29	65.4	516	1	PCT-US02-30474-1191	Sequence 1191, Ap
30	65.4	516	27	US-60-324-631-1196	Sequence 1196, Ap
31	65.4	947	1	PCT-US00-00724-73	Sequence 73, Appl
32	65.4	947	21	US-09-704-302A-73	Sequence 73, Appl
33	65.4	947	25	US-10-101-464A-73	Sequence 73, Appl
34	65.4	947	27	US-60-162-866-73	Sequence 4, Appli
35	65.4	1163	1	PCT-US02-11682-4	Sequence 4, Appli
36	65.4	1163	25	US-10-122-067-4	Sequence 2902, Ap
37	65.4	1191	19	US-09-540-236-2902	Sequence 2902, Ap
38	65.4	1191	27	US-60-128-476-4385	Sequence 4385, Ap
39	65.4	1407	17	US-09-328-352-7885	Sequence 7885, Ap
40	65.4	1499	1	PCT-US02-07787-67	Sequence 67, Appl
41	65.4	1499	1	PCT-US02-11682-2	Sequence 2, Appli
42	65.4	1499	1	PCT-US02-11760-2	Sequence 2, Appli
43	65.4	1499	24	US-10-096-534-67	Sequence 67, Appl
44	65.4	1499	25	US-10-122-067-2	Sequence 2, Appli
45	63.5	12	19	US-09-573-318-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
PCT-US01-08631-33980
; Sequence 33980, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 33980
; LENGTH: 1022
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (986)..(1003)
; OTHER INFORMATION: Zinc finger C2H2 type, domain proteins. domain identified by
; OTHER INFORMATION: eMATRIX, accession number BL00028, p-value=8.714e-09, raw score of
; OTHER INFORMATION: 16.07
; NAME/KEY: DOMAIN
; LOCATION: (472)..(1007)
; OTHER INFORMATION: Zinc finger, C2H2 type domain identified by Pfam, accession
; OTHER INFORMATION: name zf-C2H2, E-value=4.2e-28, Pfam score of 106.8

```

; NAME/KEY: misc_feature
; LOCATION: (1)...(1022)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
PCT-US01-08631-33980

```

Query Match 76.9%; Score 40; DB 1; Length 1022;
Best Local Similarity 80.0%; Pred. No. 1e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EVVPXGMSYS 11
Db 64 EVVPFGVSYS 73

```

RESULT 2
PCT-US01-08631-36185
; Sequence 36185, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/549,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 36185
; LENGTH: 1022
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (986)..(1003)
; OTHER INFORMATION: Zinc finger C2H2 type, domain identified by
; OTHER INFORMATION: eMATRIX, accession number BL00028, p-value=8.714e-09, raw score
; OTHER INFORMATION: 16.07
; NAME/KEY: DOMAIN
; LOCATION: (472)..(1007)
; OTHER INFORMATION: Zinc finger, C2H2 type domain identified by Pfam, accession
; OTHER INFORMATION: name zf-C2H2, E-value=4.2e-28, Pfam score of 106.8

```

```

; LOCATION: (1)...(1022)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
PCT-US01-08631-36185

```

Query Match 76.9%; Score 40; DB 1; Length 1022;
Best Local Similarity 80.0%; Pred. No. 1e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EVVPXGMSYS 11
Db 64 EVVPPGVSY 73

```

RESULT 3
PCT-US01-08631-38532
; Sequence 38532, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq. Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 38532

```

2.

```
; LENGTH: 3472
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-10-027-801-4

Query Match      73.1%; Score 38; DB 24; Length 3472;
Best Local Similarity 54.5%; Pred. No. 1.1e+03;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      1 BEVVPXGMSYS 11
      |:|:|:|:|:|
Db      2294 EDVIPRGISFS 2304

RESULT 6
US-10-027-806-4
; Sequence 4, Application US/10027806
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCCP.002A
; CURRENT APPLICATION NUMBER: US/10/027.806
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 3472
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-10-027-806-4

Query Match      73.1%; Score 38; DB 24; Length 3472;
Best Local Similarity 54.5%; Pred. No. 1.1e+03;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      1 BEVVPXGMSYS 11
      |:|:|:|:|:|
Db      2294 EDVIPRGISFS 2304

RESULT 7
US-10-029-120-4
; Sequence 4, Application US/10029120
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCCP.002A
; CURRENT APPLICATION NUMBER: US/10/029.120
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 3472
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-10-029-120-4

Query Match      73.1%; Score 38; DB 24; Length 3472;
Best Local Similarity 54.5%; Pred. No. 1.1e+03;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      1 BEVVPXGMSYS 11
      |:|:|:|:|:|
Db      2294 EDVIPRGISFS 2304

RESULT 8
US-10-034-623-4
; Sequence 4, Application US/10034623
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCCP.002A
; CURRENT APPLICATION NUMBER: US/10/034.623
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: 60/102,294
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 3472
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-10-034-623-4

Query Match      73.1%; Score 38; DB 24; Length 3472;
Best Local Similarity 54.5%; Pred. No. 1.1e+03;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      1 BEVVPXGMSYS 11
      |:|:|:|:|:|
Db      2294 EDVIPRGISFS 2304

RESULT 9
US-60-196-710-6207
; Sequence 6207, Application US/60196710
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000450
; CURRENT APPLICATION NUMBER: US/60/196,710
; CURRENT FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 7166
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6207
; LENGTH: 129
; TYPE: PRT
; ORGANISM: HUMAN
US-60-196-710-6207

Query Match      69.2%; Score 36; DB 27; Length 129;
Best Local Similarity 70.0%; Pred. No. 64;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 BEVVPXGMSY 10
      :|:|:|:|
Db      103 KEVPTGHSY 112

RESULT 10
US-60-177-646-2588
; Sequence 2588, Application US/60177646
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000210
; CURRENT APPLICATION NUMBER: US/60/177,646
; CURRENT FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 4226
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2588
; LENGTH: 140
; TYPE: PRT
; ORGANISM: HUMAN
US-60-177-646-2588

Query Match          69.2%; Score 36; DB 27; Length 140;
Best Local Similarity 70.0%; Pred. No. 70;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 BEVVPXGMSY 10
Db 97 KEVVTGHSY 106

RESULT 11
US-10-029-386-32486
; Sequence 32486, Application US/10029386
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hazzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 32486
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AF002994.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
; OTHER INFORMATION: SWISSPROT HIT: P43363, EVALUE 5.00e-46
US-10-029-386-32486

Query Match          69.2%; Score 36; DB 24; Length 153;
Best Local Similarity 70.0%; Pred. No. 78;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 BEVVPXGMSY 10
Db 15 KEVVTGHSY 24

RESULT 12
US-09-134-000-3738
; Sequence 3738, Application US/09134000A
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS
; FILE REFERENCE: GTC-005
; CURRENT APPLICATION NUMBER: US/09/134,000A
; CURRENT FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 6810
; SEQ ID NO 3738
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (327), (328)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unkno
US-09-134-000-3738
```

```
Query Match          69.2%; Score 36; DB 15; Length 382;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VVPXGMSYS 11
Db 332 LIPEGMSYS 340

RESULT 13
US-60-389-987-2598
; Sequence 2598, Application US/60389987
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465P2
; CURRENT APPLICATION NUMBER: US/60/389,987
; CURRENT FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 3025
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2598
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-389-987-2598

Query Match          69.2%; Score 36; DB 27; Length 401;
Best Local Similarity 70.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 BEVVPXGMSY 10
Db 263 KEVVTGHSY 272

RESULT 14
US-60-412-418-2598
; Sequence 2598, Application US/60412418
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465P3
; CURRENT APPLICATION NUMBER: US/60/412,418
; CURRENT FILING DATE: 2002-09-20
; NUMBER OF SEQ ID NOS: 3025
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2598
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-412-418-2598

Query Match          69.2%; Score 36; DB 27; Length 401;
Best Local Similarity 70.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 BEVVPXGMSY 10
Db 263 KEVVTGHSY 272
```

Db 263 KEVVTGHSY 272

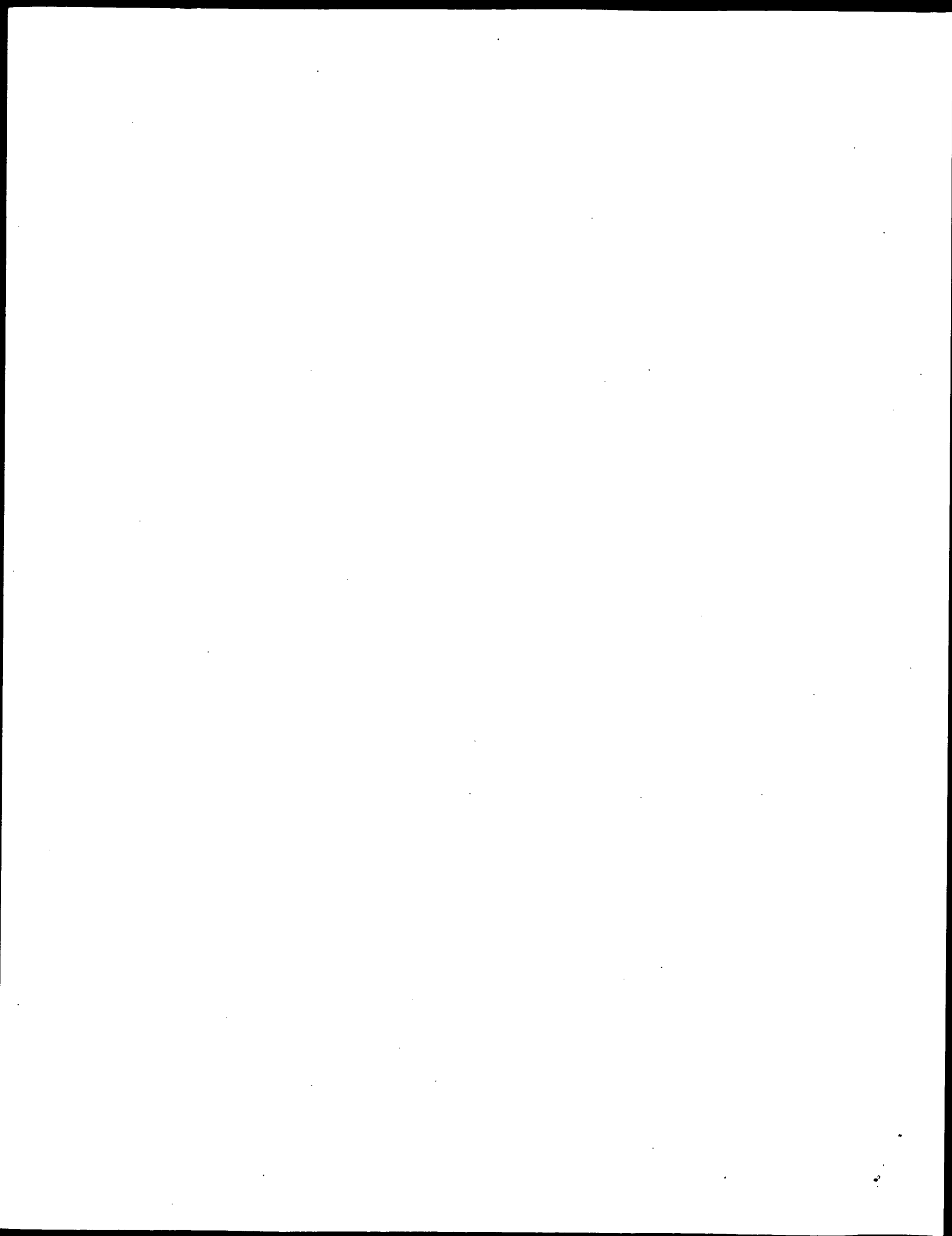
RESULT 15

US-10-219-999-58270
; Sequence 58270, Application US/10219999
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: cDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)C
; CURRENT APPLICATION NUMBER: US/10/219,999
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 60/324,109
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/312,544
; PRIOR FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 63520
; SEQ ID NO 58270
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Zea mays
US-10-219-999-58270

Query Match 67.3%; Score 35; DB 26; Length 160;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10
Db 23 ELLPVGMSY 31

Search completed: June 4, 2003, 13:27:19
Job time : 185 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:13:35 ; Search time 14.25 Seconds
(without alignments)
79.694 Million cell updates/sec

Title: AUDET-909-1
Perfect score: 52
Sequence: 1 eevvpxgmsys 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	73.1	3472	9	US-10-027-806-4
2	38	73.1	3472	9	US-10-034-623-4
3	38	73.1	3472	9	US-10-027-801-4
4	34	65.4	947	9	US-10-101-464A-73
5	33	63.5	426	9	US-10-214-766-43
6	33	63.5	478	9	US-09-924-340-108
7	33	63.5	478	9	US-09-992-600A-108
8	33	63.5	478	9	US-09-746-783-184
9	33	63.5	478	9	US-10-000-489-108
10	33	63.5	478	9	US-10-000-985-108
11	33	63.5	653	9	US-09-820-843A-26
12	32	61.5	10	10	US-09-947-387-66
13	32	61.5	254	10	US-09-778-927A-53
14	32	61.5	289	9	US-10-024-066-2
15	32	61.5	289	9	US-10-024-066-4
16	32	61.5	289	10	US-09-919-497-54
17	32	61.5	295	10	US-09-925-300-1061
18	32	61.5	529	10	US-09-923-304-4
19	32	61.5	691	9	US-10-101-921-4

20	32	61.5	691	10	US-09-925-731-2	Sequence 2, Appli
21	32	61.5	1377	10	US-09-815-242-10384	Sequence 10384, A
22	32	61.5	2799	9	US-10-151-736-4	Sequence 4, Appli
23	31	59.6	53	9	US-10-092-154-878	Sequence 878, App
24	31	59.6	53	10	US-09-764-847-878	Sequence 878, App
25	31	59.6	59	10	US-09-948-080-14	Sequence 14, Appli
26	31	59.6	161	9	US-09-738-626-5124	Sequence 5124, Ap
27	31	59.6	163	9	US-10-117-846-20	Sequence 20, Appli
28	31	59.6	192	9	US-09-986-480-171	Sequence 171, App
29	31	59.6	198	10	US-09-731-872-334	Sequence 334, App
30	31	59.6	223	9	US-09-738-626-6349	Sequence 6349, Ap
31	31	59.6	239	10	US-09-815-242-10697	Sequence 10697, A
32	31	59.6	381	9	US-09-975-139-5	Sequence 5, Appli
33	31	59.6	702	9	US-10-280-403-2	Sequence 2, Appli
34	31	59.6	702	10	US-09-907-479-2	Sequence 2, Appli
35	31	59.6	763	9	US-09-738-626-4454	Sequence 4454, Ap
36	31	59.6	1053	10	US-09-815-242-5136	Sequence 5136, Ap
37	31	59.6	1407	10	US-09-815-242-10439	Sequence 10439, A
38	31	59.6	1426	10	US-09-912-020-340	Sequence 340, App
39	30	57.7	7	9	US-09-909-062-1	Sequence 1, Appli
40	30	57.7	7	9	US-09-909-062-9	Sequence 9, Appli
41	30	57.7	7	9	US-09-909-062-130	Sequence 130, App
42	30	57.7	121	9	US-09-852-797-68	Sequence 68, Appli
43	30	57.7	121	9	US-09-852-797-85	Sequence 85, Appli
44	30	57.7	121	10	US-09-853-161-68	Sequence 68, Appli
45	30	57.7	121	10	US-09-853-161-85	Sequence 85, Appli

ALIGNMENTS

RESULT 1

US-10-027-806-4
; Sequence 4, Application US/10027806
; Patent No. US20020160476A1
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCOIP.002A
; CURRENT APPLICATION NUMBER: US/10/027,806
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 3472
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-10-027-806-4

Query Match 73.1%; Score 38; DB 9; Length 3472;
Best Local Similarity 54.5%; Pred.No. 1.2e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11

Db 2294 EDVIPRGISFS 2304

RESULT 2

US-10-034-623-4
; Sequence 4, Application US/10034623
; Publication No. US20020198365A1
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCOIP.002A
; CURRENT APPLICATION NUMBER: US/10/034,623

audet-909-1.rapb

Wed Jun 4 16:50:26 2003

```
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: 60/102,294
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 3472
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-10-034-623-4

Query Match      73.1%; Score 38; DB 9; Length 3472;
Best Local Similarity 54.5%; Pred. No. 1.2e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVVVPXGMSYS 11
|:|:|:|:|:|
Db 2294 EDVIPRGISFS 2304

RESULT 3
US-10-027-801-4
; Sequence 4, Application US/10027801
; Publication No. US20030054364A1
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCCP.002A
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: 1999-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 3472
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-10-027-801-4

Query Match      73.1%; Score 38; DB 9; Length 3472;
Best Local Similarity 54.5%; Pred. No. 1.2e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVVVPXGMSYS 11
|:|:|:|:|:|
Db 2294 EDVIPRGISFS 2304

RESULT 4
US-10-101-464A-73
; Sequence 73, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000.1020C2
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
```

```
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-101-464A-73

Query Match      65.4%; Score 34; DB 9; Length 947;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11
|:|:|:|:|
Db 686 VMPSGISYS 694

RESULT 5
US-10-214-766-43
; Sequence 43, Application US/10214766
; Publication No. US20030084473A1
; GENERAL INFORMATION:
; APPLICANT: Gocal, Greg
; TITLE OF INVENTION: NON-TRANSGENIC HERBICIDE RESISTANT PLANTS
; FILE REFERENCE: CA1138
; CURRENT APPLICATION NUMBER: US/10/214,766
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,734
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 43
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-214-766-43

Query Match      63.5%; Score 33; DB 9; Length 426;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVVVPXGMSYS 10
|:|:|:|:|
Db 223 EFVIPAGQSY 232

RESULT 6
US-09-924-340-108
; Sequence 108, Application US/09924340
; Publication No. US20030027248A1
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US2.REG
; CURRENT APPLICATION NUMBER: US/09/924,340
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 108
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-924-340-108
```


Query Match 63.5%; Score 33; DB 9; Length 478;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11
|||
DB 239 EVAPAGASYN 248

RESULT 7

US-09-992-600A-108
; Sequence 108, Application US/09992600A
; Publication No. US20030027161A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US4.DIV
; CURRENT APPLICATION NUMBER: US/09/992,600A
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: JPatent
; SEQ ID NO 108
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-992-600A-108

Query Match 63.5%; Score 33; DB 9; Length 478;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11
|||
DB 239 EVAPAGASYN 248

RESULT 8

US-09-746-783-184
; Sequence 184, Application US/09746783
; Publication No. US20030044935A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Racie, Lisa A.
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Fechtel, Kim
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 231
; CORRESPONDENCE ADDRESS:
; ADDRESS: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: U.S.A.

ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/746,783
FILING DATE: 21-Dec-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Milasincic, Debra J.
REGISTRATION NUMBER: 46,931
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 184:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 184:
US-09-746-783-184

Query Match 63.5%; Score 33; DB 9; Length 478;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11
|||
DB 239 EVAPAGASYN 248

RESULT 9

US-10-000-489-108
; Sequence 108, Application US/10000489
; Publication No. US20030092011A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US6.DIV
; CURRENT APPLICATION NUMBER: US/10/000,489
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 108
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-000-489-108

Query Match 63.5%; Score 33; DB 9; Length 478;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11
|||
DB 239 EVAPAGASYN 248

RESULT 10

US-10-000-986-108
; Sequence 108, Application US/10000986
; Publication No. US20030096247A1
; GENERAL INFORMATION:
; APPLICANT: Benjanin, Stephanie
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US9.DIV
; CURRENT APPLICATION NUMBER: US/10/000,986
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 108
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-000-986-108

Query Match 63.5%; Score 33; DB 9; Length 478;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11
Db 239 EVAPAGASYN 248

RESULT 11

US-09-820-843A-26
; Sequence 26, Application US/09820843A
; Publication No. US20030039963A1
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEIN
; FILE REFERENCE: Q63915
; CURRENT APPLICATION NUMBER: US/09/820,843A
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Vibrio cholerae
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: iron(III) ABC transporter, permease protein
; NAME/KEY: misc feature
; OTHER INFORMATION: gi|9654609
US-09-820-843A-26

Query Match 63.5%; Score 33; DB 9; Length 653;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVVPXGMS 9
Db 300 EVVPSGIT 308

RESULT 12

US-09-947-387-66
; Sequence 66, Application US/09947387
; Patent No. US20020150885A1
; GENERAL INFORMATION:
; APPLICANT: Weber, Eckard
; APPLICANT: Cai, Sui Xiong
; APPLICANT: Keana, John F.W.
; APPLICANT: Drewe, John A.
; APPLICANT: Zhang, Han-Zhong
; TITLE OF INVENTION: No. US20020150885A1 Fluorogenic or Fluorescent Reporter Molecule
; TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence
; TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the
; TITLE OF INVENTION: Use Thereof
; FILE REFERENCE: 1735.0290005
; CURRENT APPLICATION NUMBER: US/09/947,387
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/061,582
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: US 60/145,746
; PRIOR FILING DATE: 1998-03-03
; PRIOR APPLICATION NUMBER: US 09/168,888
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 66
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-947-387-66

Query Match 61.5%; Score 32; DB 10; Length 10;
Best Local Similarity 50.0%; Pred. No. 3.5;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
Db 1 DDIVPCMSY 10

RESULT 13

US-09-778-927A-53
; Sequence 53, Application US/09778927A
; Patent No. US20020068342A1
; GENERAL INFORMATION:
; APPLICANT: KHOSRAVI, Rami et al.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID AND AMINO ACID SEQUENCES AND NOVEL
; TITLE OF INVENTION: VARIANTS OF ALTERNATIVE SPLICING
; FILE REFERENCE: 2786-0160P
; CURRENT APPLICATION NUMBER: US/09/778,927A
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: IL 134453
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: IL135341
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(254)
; OTHER INFORMATION: Xaa = any amino acid, unknown, or other
US-09-778-927A-53

Query Match 61.5%; Score 32; DB 10; Length 254;

Search completed: June 4, 2003, 13:30:38
Job time : 15.25 secs

Best Local Similarity 60.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10
||| | | |
Db 74 EEVFPFLAMNY 83

RESULT 14

US-10-024-066-2
; Sequence 2, Application US/10024066
; Patent No. US20020166134A1
; GENERAL INFORMATION:
; APPLICANT: Field, Loren J.
; APPLICANT: Pasumarthi, Kishore Babu S.
; TITLE OF INVENTION: CARDIOMYOCYTES WITH ENHANCED PROLIFERATIVE POTENTIAL,
; TITLE OF INVENTION: AND METHODS FOR PREPARING AND USING SAME
; FILE REFERENCE: 7037-450
; CURRENT APPLICATION NUMBER: US/10/024,066
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/139,942
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: PCT/US00/16827
; PRIOR FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-024-066-2

Query Match 61.5%; Score 32; DB 9; Length 289;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

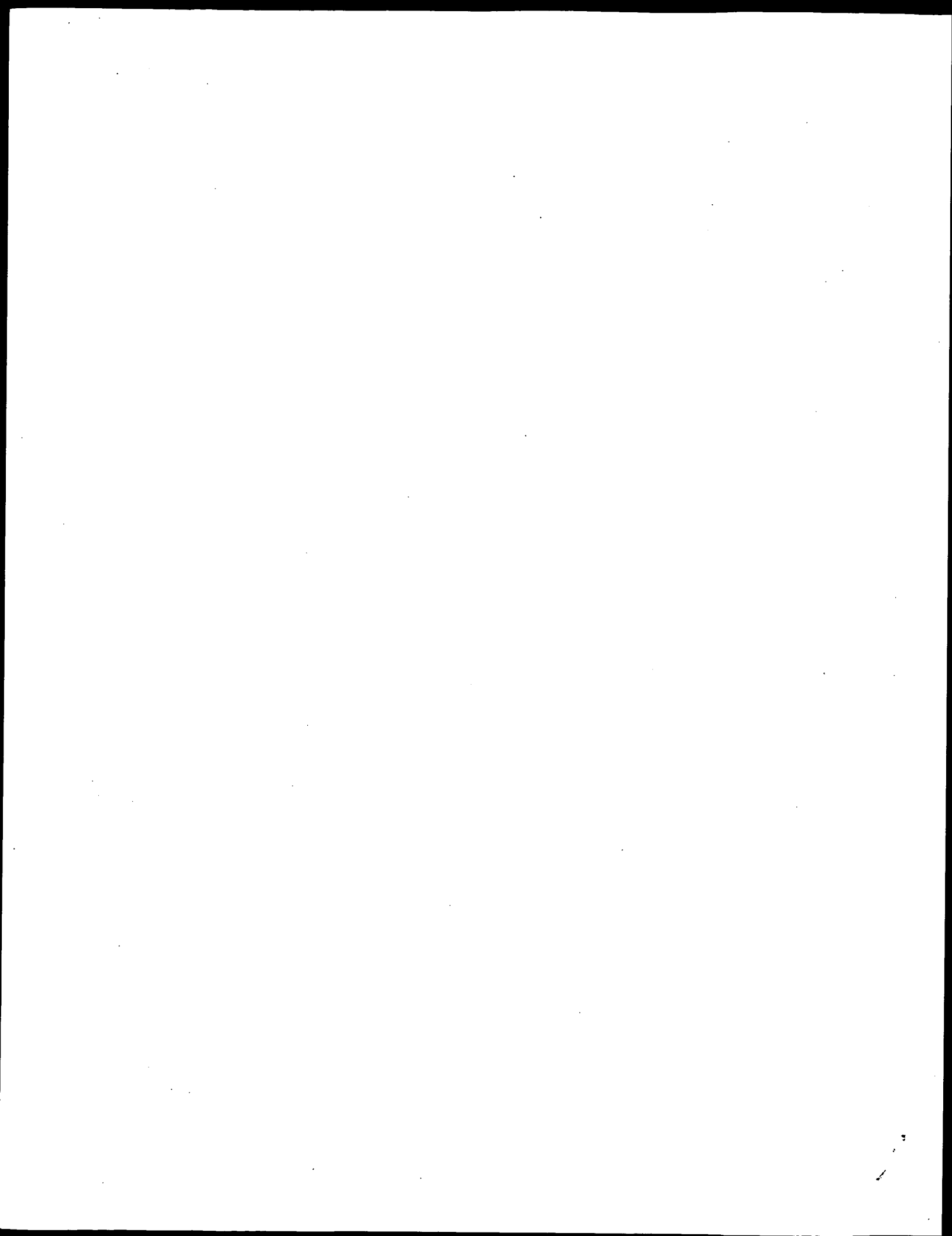
Qy 1 EEVVPXGMSY 10
||| | | |
Db 73 EEVFPFLAMNY 82

RESULT 15

US-10-024-066-4
; Sequence 4, Application US/10024066
; Patent No. US20020166134A1
; GENERAL INFORMATION:
; APPLICANT: Field, Loren J.
; APPLICANT: Pasumarthi, Kishore Babu S.
; TITLE OF INVENTION: CARDIOMYOCYTES WITH ENHANCED PROLIFERATIVE POTENTIAL,
; TITLE OF INVENTION: AND METHODS FOR PREPARING AND USING SAME
; FILE REFERENCE: 7037-450
; CURRENT APPLICATION NUMBER: US/10/024,066
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/139,942
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: PCT/US00/16827
; PRIOR FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-024-066-4

Query Match 61.5%; Score 32; DB 9; Length 289;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10
||| | | |
Db 74 EEVFPFLAMNY 83



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:09:19 ; Search time 9.75 Seconds
(without alignments)

33.195 Million cell updates/sec

Title: AUDET-909-1

Perfect score: 52

Sequence: 1 eevvpkgmsys 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	65.4	947	4	US-09-228-986-73
2	33	63.5	45	2	US-08-637-759B-236
3	33	63.5	45	3	US-08-871-355A-236
4	33	63.5	45	4	US-09-201-945-236
5	33	63.5	65	6	5177197-51
6	33	63.5	410	6	5177197-1
7	33	63.5	1394	6	5177197-30
8	32	61.5	10	4	US-09-357-952-66
9	32	61.5	10	4	US-09-521-650-66
10	32	61.5	10	4	US-09-168-888-66
11	32	61.5	102	2	US-08-580-988A-23
12	32	61.5	152	2	US-08-460-694-4
13	32	61.5	152	3	US-08-460-744-4
14	32	61.5	152	3	US-07-667-711B-4
15	32	61.5	173	1	US-08-193-977-7
16	32	61.5	189	2	US-08-464-517-21
17	32	61.5	189	2	US-08-246-361A-21
18	32	61.5	189	3	US-08-463-772-21
19	32	61.5	189	5	PCT-US93-05000-21
20	32	61.5	236	2	US-08-464-517-22
21	32	61.5	236	2	US-08-246-361A-22
22	32	61.5	236	3	US-08-463-772-22
23	32	61.5	236	5	PCT-US93-05000-22
24	32	61.5	280	2	US-08-464-517-6
25	32	61.5	280	3	US-08-463-772-6
26	32	61.5	289	2	US-08-246-361A-4
27	32	61.5	289	5	PCT-US93-05000-4

28 32 61.5 291 5 PCT-US93-05000-6 Sequence 6, Appli
29 32 61.5 292 2 US-08-464-517-23 Sequence 23, Appli
30 32 61.5 292 2 US-08-246-361A-6 Sequence 6, Appli
31 32 61.5 292 2 US-08-246-361A-23 Sequence 23, Appli
32 61.5 292 3 US-08-463-772-23 Sequence 23, Appli
33 32 61.5 292 5 PCT-US93-05000-23 Sequence 23, Appli
34 32 61.5 295 1 US-07-947-120-8 Sequence 8, Appli
35 32 61.5 295 1 US-08-472-893A-8 Sequence 8, Appli
36 32 61.5 295 2 US-08-460-694-2 Sequence 2, Appli
37 32 61.5 295 2 US-08-464-517-19 Sequence 19, Appli
38 32 61.5 295 2 US-08-464-517-20 Sequence 19, Appli
39 32 61.5 295 2 US-08-246-361A-19 Sequence 19, Appli
40 32 61.5 295 2 US-08-246-361A-20 Sequence 20, Appli
41 32 61.5 295 3 US-08-463-772-19 Sequence 19, Appli
42 32 61.5 295 3 US-08-463-772-20 Sequence 20, Appli
43 32 61.5 295 3 US-08-460-744-2 Sequence 2, Appli
44 32 61.5 295 3 US-07-667-711B-2 Sequence 2, Appli
45 32 61.5 295 3 US-08-947-492-8 Sequence 8, Appli

ALIGNMENTS

RESULT 1

US-09-228-986-73
; Sequence 73, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 73
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-228-986-73

Query Match 65.4%; Score 34; DB 4; Length 947;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11
|:|:|:|
Db 686 VMPGSGISYS 694

RESULT 2

US-08-637-759B-236
; Sequence 236, Application US/08637759B
; Patent No. 5876931
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30

```

,
, CURRENT APPLICATION DATA:
, APPLICATION NUMBER: US/08/637,759B
, FILING DATE: 03-MAY-1996
, CLASSIFICATION: 435
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: PCT/GB95/02875
, FILING DATE: 11-DEC-1995
, CLASSIFICATION: 435
, ATTORNEY/AGENT INFORMATION:
, NAME: Pabst, Patrea L.
, REGISTRATION NUMBER: 31,284
, REFERENCE/DOCKET NUMBER: RPNs 101
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: (404) 873-8794
, TELEFAX: (404) 873-8795
, INFORMATION FOR SEQ ID NO: 236:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 45 amino acids
, TYPE: amino acid
, STRANDEDNESS: single
, TOPOLOGY: linear
, MOLECULE TYPE: protein
, HYPOTHETICAL: NO
, US-08-637-759B-236

```

Query Match 63.5%; Score 33; DB 2; Length 45;
Best Local Similarity 60.0%;
Matches 6; Conservative 1; Mismatches 3; Indels

Qy 1 EEVVPXGMSY 10
|||:
Db 1 EEISPLGWSY 10

RESULT 3
US-08-871-355A-236
Sequence 236, Application US/08871355A
Patent No. 6015669
GENERAL INFORMATION:
APPLICANT: David William Holden
TITLE OF INVENTION: Identification of Genes
NUMBER OF SEQUENCES: 501
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #4
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/871,355A
FILING DATE: 09-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02875
FILING DATE: 11-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RWS 101 CON
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 236:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids

```

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-871-355A-236

```

Query Match 63.5%; Score 33; DB 3; Length 45;
Best Local Similarity 60.0%; Pred. No. 5.1;
Matches 6; Conservative 1; Mismatches 3; Indels

Qy 1 EEVVPXGMSY 10
||: || ||
Db 1 EEISPLGWSY 10

```

1  RESULT 4
2  US-09-201-945-236
3  ; Sequence 236, Application US/09201945
4  ; Patent No. 6342215
5  ; GENERAL INFORMATION:
6  ; APPLICANT: David William Holden
7  ; TITLE OF INVENTION: Identification of Genes
8  ; NUMBER OF SEQUENCES: 501
9  ; CORRESPONDENCE ADDRESS:
10 ; ADDRESSEE: Patrea L. Pabst
11 ; STREET: 2800 One Atlantic Center
12 ; STREET: 1201 West Peachtree Street
13 ; CITY: Atlanta
14 ; STATE: Georgia
15 ; COUNTRY: USA
16 ; ZIP: 30309-3450
17 ; COMPUTER READABLE FORM:
18 ; MEDIUM TYPE: Floppy disk
19 ; COMPUTER: IBM PC compatible
20 ; OPERATING SYSTEM: PC-DOS/MS-DOS
21 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
22 ; CURRENT APPLICATION DATA:
23 ; APPLICATION NUMBER: US/09/201.945
24 ; FILING DATE:
25 ; CLASSIFICATION:
26 ; PRIOR APPLICATION DATA:
27 ; APPLICATION NUMBER: 08/637,759
28 ; FILING DATE:
29 ; CLASSIFICATION:
30 ; ATTORNEY/AGENT INFORMATION:
31 ; NAME: Pabst, Patrea L.
32 ; REGISTRATION NUMBER: 31,284
33 ; REFERENCE/DOCKET NUMBER: RPS 101
34 ; TELECOMMUNICATION INFORMATION:
35 ; TELEPHONE: (404) 873-8794
36 ; TELEFAX: (404) 873-8795
37 ; INFORMATION FOR SEQ ID NO: 236:
38 ; SEQUENCE CHARACTERISTICS:
39 ; LENGTH: 45 amino acids
40 ; TYPE: amino acid
41 ; STRANDEDNESS: single
42 ; TOPOLOGY: linear
43 ; MOLECULE TYPE: protein
44 ; HYPOTHETICAL: NO
45 ;
46 US-09-201-945-236

```

Query Match 63.5%; Score 33; DB 4; Length 45;
Best Local Similarity 60.0%; Pred. No. 5.1;
Matches 6; Conservative 1; Mismatches 3; Indels

Qy 1 EEVVPXGMSY 10
|||: |||
Db 1 EEISPLGWSY 10

RESULT 5
5177197-51

Patent No. 5177197
; APPLICANT: KANZAKI, TETSUTO; OLOFSSON, ANDERS; MOREN, ANITA;
; WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAES-
; SON-WELSH, LENA; HELDIN, CARL-HENRIK
; TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING
; HUMAN TRANSFORMING GROWTH FACTOR-BETAL-BINDING PROTEIN
; NUMBER OF SEQUENCES: 53
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/487,343
; FILING DATE: 27-FEB-1990
; SEQ ID NO: 51:
; LENGTH: 65
5177197-51

Query Match 63.5%; Score 33; DB 6; Length 65;
Best Local Similarity 45.5%; Pred. No. 7.8;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 BEVVPXGMSYS 11
:|:|:|:|:
Db 52 KEICPGMGYT 62

RESULT 6

Patent No. 5177197
; APPLICANT: KANZAKI, TETSUTO; OLOFSSON, ANDERS; MOREN, ANITA;
; WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAES-
; SON-WELSH, LENA; HELDIN, CARL-HENRIK
; TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING
; HUMAN TRANSFORMING GROWTH FACTOR-BETAL-BINDING PROTEIN
; NUMBER OF SEQUENCES: 53
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/487,343
; FILING DATE: 27-FEB-1990
; SEQ ID NO: 1:
; LENGTH: 410
5177197-1

Query Match 63.5%; Score 33; DB 6; Length 410;
Best Local Similarity 45.5%; Pred. No. 63;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 BEVVPXGMSYS 11
:|:|:|:|:
Db 399 KEICPGMGYT 409

RESULT 7

Patent No. 5177197
; APPLICANT: KANZAKI, TETSUTO; OLOFSSON, ANDERS; MOREN, ANITA;
; WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAES-
; SON-WELSH, LENA; HELDIN, CARL-HENRIK
; TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING
; HUMAN TRANSFORMING GROWTH FACTOR-BETAL-BINDING PROTEIN
; NUMBER OF SEQUENCES: 53
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/487,343
; FILING DATE: 27-FEB-1990
; SEQ ID NO: 30:
; LENGTH: 1394
5177197-30

Query Match 63.5%; Score 33; DB 6; Length 1394;
Best Local Similarity 45.5%; Pred. No. 2.6e-02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 BEVVPXGMSYS 11
:|:|:|:|:
Db 399 KEICPGMGYT 409

RESULT 8

US-09-357-952-66
; Sequence 66, Application US/09357952
; Patent No. 6248904
; GENERAL INFORMATION:
; APPLICANT: Zhang, Han-Zhong
; APPLICANT: Cai, Sui Xiong
; APPLICANT: Drewe, John A.
; APPLICANT: Yang, Wu
; TITLE OF INVENTION: No. 6248904el Fluorescence Dyes and Their Applications for Whole-
; TITLE OF INVENTION: Fluorescence Screening Assays for Caspases, Peptidases, Proteases
; TITLE OF INVENTION: Other Enzymes and the Use Thereof
; FILE REFERENCE: 1735.0030001
; CURRENT APPLICATION NUMBER: US/09/357,952
; CURRENT FILING DATE: 1999-07-21
; EARLIER APPLICATION NUMBER: US 60/093,642
; EARLIER FILING DATE: 21-JUL-1998
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 66
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-357-952-66

Query Match 61.5%; Score 32; DB 4; Length 10;
Best Local Similarity 50.0%; Pred. No. 1.5;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 BEVVPXGMSY 10
:|:|:|:|:
Db 1 DDIVPCMSY 10

RESULT 9

US-09-521-650-66
; Sequence 66, Application US/09521650
; Patent No. 6335429
; GENERAL INFORMATION:
; APPLICANT: Weber, Eckard
; APPLICANT: Cai, Sui Xiong
; APPLICANT: Keana, John F.W.
; APPLICANT: Drewe, John A.
; APPLICANT: Zhang, Han-Zhong
; TITLE OF INVENTION: No. 6335429el Fluorogenic or Fluorescent Reporter Molecules and
; TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence
; TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the
; TITLE OF INVENTION: Use Thereof
; FILE REFERENCE: 1735.0290002
; CURRENT APPLICATION NUMBER: US/09/521,650
; CURRENT FILING DATE: 2000-03-08
; EARLIER APPLICATION NUMBER: 09/168,888
; EARLIER FILING DATE: 1998-10-09
; EARLIER APPLICATION NUMBER: US 60/061,582
; EARLIER FILING DATE: 1997-10-10
; EARLIER APPLICATION NUMBER: US 09/033,661
; EARLIER FILING DATE: 1998-03-03
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 66
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-521-650-66

Query Match 61.5%; Score 32; DB 4; Length 10;
Best Local Similarity 50.0%; Pred. No. 1.5;

Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10
: : : : :
Db 1 DDIVPCMSY 10

RESULT 10
US-09-168-888-66
; Sequence 66, Application US/09168888
; Patent No. 6342611
; GENERAL INFORMATION:
; APPLICANT: Weber, Eckard
; APPLICANT: Cai, Sui Xiong
; APPLICANT: Keana, John F.W.
; APPLICANT: Drewe, John A.
; APPLICANT: Zhang, Han-Zhong
; TITLE OF INVENTION: No. 6342611el Fluorogenic or Fluorescent Reporter Molecules and
; TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence
; TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the
; TITLE OF INVENTION: Use Thereof
; FILE REFERENCE: 1735.0290002
; CURRENT APPLICATION NUMBER: US/09/168,888
; EARLIER FILING DATE: 1998-10-09
; EARLIER FILING DATE: 1997-10-10
; EARLIER APPLICATION NUMBER: US 09/033,661
; EARLIER FILING DATE: 1998-03-03
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 66
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-168-888-66

Query Match 61.5%; Score 32; DB 4; Length 10;
Best Local Similarity 50.0%; Pred. No. 1.5;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10
: : : : :
Db 1 DDIVPCMSY 10

RESULT 11
US-08-580-988A-23
; Sequence 23, Application US/08580988A
; Patent No. 5856161
; GENERAL INFORMATION:
; APPLICANT: Aggarwal et al.
; TITLE OF INVENTION: Tumor Necrosis Factor
; TITLE OF INVENTION: Receptor-I-Associated Protein Kinase And Methods
; TITLE OF INVENTION: For Its Use
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. Benjamin A. Adler
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/580,988A
; FILING DATE: January 3, 1996

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benjamin Aaron Adler, Ph.D., J.D.
REGISTRATION NUMBER: 35,423
REFERENCE/DOCKET NUMBER: D5721CIP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713-777-2321
TELEFAX: 713-777-6908
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 102 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: protein
HYPOTHETICAL: no
ANTI-SENSE: no
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
US-08-580-988A-23

Query Match 61.5%; Score 32; DB 2; Length 102;
Best Local Similarity 60.0%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10
: : : : :
Db 24 EEVFPPLANNY 33

RESULT 12
US-08-460-694-4
; Sequence 4, Application US/08460694
; Patent No. 5858655
; GENERAL INFORMATION:
; APPLICANT: Arnold, Andrew
; TITLE OF INVENTION: PRAD1 Cyclin and its cDNA
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,694
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McConathy, Evelyn H.
; REGISTRATION NUMBER: 35,279
; REFERENCE/DOCKET NUMBER: 0609.4070002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 152 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-460-694-4

Query Match 61.5%; Score 32; DB 2; Length 152;
 Best Local Similarity 60.0%; Pred. No. 33;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
 ||||| |:
 Db 20 EEVFPPLAMNY 29

RESULT 13

US-08-460-744-4
 ; Sequence 4, Application US/08460744
 ; Patent No. 6107541
 ; GENERAL INFORMATION:
 ; APPLICANT: Arnold, Andrew
 ; TITLE OF INVENTION: Pradi Cyclin and its cDNA
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
 ; STREET: 1100 New York Avenue, N.W., Suite 600
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20005

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/460,744
 FILING DATE: 02-JUN-1995
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 NAME: McConathy, Evelyn H.
 REGISTRATION NUMBER: 35,279
 REFERENCE/DOCKET NUMBER: 0609.4070005
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-371-2600
 TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 152 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-460-744-4

Query Match 61.5%; Score 32; DB 3; Length 152;
 Best Local Similarity 60.0%; Pred. No. 33;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
 ||||| |:
 Db 20 EEVFPPLAMNY 29

RESULT 14

US-07-667-711B-4
 ; Sequence 4, Application US/07667711B
 ; Patent No. 6110700
 ; GENERAL INFORMATION:
 ; APPLICANT: ARNOLD, ANDREW
 ; TITLE OF INVENTION: Pradi Cyclin and Its cDNA
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
 ; STREET: 1100 NEW YORK AVE., NW, SUITE 600
 ; CITY: WASHINGTON
 ; STATE: DC
 ; COUNTRY: USA

ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/667,711B
 FILING DATE: 11-MAR-1991
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: MCPHAIL, DONALD R.
 REGISTRATION NUMBER: 35,811
 REFERENCE/DOCKET NUMBER: 0609.4070000
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 152 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: not relevant
 MOLECULE TYPE: peptide
 US-07-667-711B-4

Query Match 61.5%; Score 32; DB 3; Length 152;
 Best Local Similarity 60.0%; Pred. No. 33;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
 ||||| |:
 Db 20 EEVFPPLAMNY 29

RESULT 15

US-08-193-977-7
 ; Sequence 7, Application US/08193977
 ; Patent No. 5625031
 ; GENERAL INFORMATION:
 ; APPLICANT: WEBSTER, KEVIN R.
 ; APPLICANT: COLEMAN, KEVIN G.
 ; TITLE OF INVENTION: PEPTIDE INHIBITORS OF THE P33CDK2 AND
 ; P34CDC2 CELL CYCLE REGULATORY KINASES AND HUMAN
 ; TITLE OF INVENTION: PAPILLOMAVIRUS E7 ONCOPROTEIN
 ; NUMBER OF SEQUENCES: 34
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: REED & ROBINS
 ; STREET: 635 BRYANT STREET
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: UNITED STATES OF AMERICA
 ; ZIP: 94301

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/193,977
 FILING DATE: 08-FEB-1994
 CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
 NAME: ROBINS, ROBERTA L.
 REGISTRATION NUMBER: 33,208
 REFERENCE/DOCKET NUMBER: S998-0016
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 617-8999
 TELEFAX: (415) 327-3231
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 173 amino acids
 TYPE: amino acid

Search completed: June 4, 2003, 13:14:58
Job time : 9.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:04:09 ; Search time 28.25 Seconds
(without alignments)
51.885 Million cell updates/sec

Title: AUDET-909-2

Perfect score: 56

Sequence: 1 eevvpxgmhys 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

1:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	96.4	11	23	ABB80523 Hepatitis C virus
2	54	96.4	11	23	ABB80527 Hepatitis C virus
3	54	96.4	11	23	ABB80558 Hepatitis C virus
4	54	96.4	11	23	ABB80560 Hepatitis C virus
5	49	87.5	11	23	ABB80537 Hepatitis C virus
6	49	87.5	11	23	ABB80541 Hepatitis C virus
7	48	85.7	11	23	ABB80546 Hepatitis C virus
8	48	85.7	11	23	ABB80550 Hepatitis C virus
9	48	85.7	11	23	ABB80554 Hepatitis C virus
10	48	85.7	11	23	ABB80555 Hepatitis C virus

11	46	82.1	11	23	ABB80531 Hepatitis C virus
12	46	82.1	11	23	ABB80532 Hepatitis C virus
13	45	80.4	11	23	ABB80521 Hepatitis C virus
14	45	80.4	11	23	ABB80522 Hepatitis C virus
15	45	80.4	11	23	ABB80524 Hepatitis C virus
16	45	80.4	11	23	ABB80525 Hepatitis C virus
17	45	80.4	11	23	ABB80526 Hepatitis C virus
18	45	80.4	11	23	ABB80528 Hepatitis C virus
19	45	80.4	11	23	ABB80529 Hepatitis C virus
20	45	80.4	11	23	ABB80559 Hepatitis C virus
21	45	80.4	11	23	ABB80561 Hepatitis C virus
22	45	80.4	11	23	ABB80562 Hepatitis C virus
23	45	80.4	11	23	ABB80563 Hepatitis C virus
24	45	80.4	11	23	ABB80564 Hepatitis C virus
25	45	80.4	11	23	ABB80565 Hepatitis C virus
26	45	80.4	11	23	ABB80566 Hepatitis C virus
27	45	80.4	11	23	ABB80567 Hepatitis C virus
28	45	80.4	11	23	ABB80568 Hepatitis C virus
29	40	71.4	11	23	ABB80535 Hepatitis C virus
30	40	71.4	11	23	ABB80536 Hepatitis C virus
31	40	71.4	11	23	ABB80538 Hepatitis C virus
32	40	71.4	11	23	ABB80539 Hepatitis C virus
33	40	71.4	11	23	ABB80540 Hepatitis C virus
34	40	71.4	11	23	ABB80542 Hepatitis C virus
35	40	71.4	11	23	ABB80543 Hepatitis C virus
36	39	69.6	11	23	ABB80544 Hepatitis C virus
37	39	69.6	11	23	ABB80545 Hepatitis C virus
38	39	69.6	11	23	ABB80547 Hepatitis C virus
39	39	69.6	11	23	ABB80548 Hepatitis C virus
40	39	69.6	11	23	ABB80549 Hepatitis C virus
41	39	69.6	11	23	ABB80551 Hepatitis C virus
42	39	69.6	11	23	ABB80552 Hepatitis C virus
43	39	69.6	11	23	ABB80553 Hepatitis C virus
44	39	69.6	11	23	ABB80556 Hepatitis C virus
45	39	69.6	11	23	ABB80557 Hepatitis C virus

ALIGNMENTS

RESULT 1
ABB80523
ID ABB80523 standard; peptide; 11 AA.
XX ABB80523;
XX AC
XX 08-OCT-2002 (first entry)
DT Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #3.
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide; virucide.
XX Synthetic.
XX Key Location/Qualifiers
FH Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with residue 7"
FT Misc-difference 9 /note= "D-form residue"
FT Modified-site 11 /note= "C-terminal amide"
XX WO200208251-A2.
PN 31-JAN-2002.
PD 19-JUL-2001; 2001WO-US23169.
XX 21-JUL-2000; 2000US-220101P.

XX PA (CORV-) CORVAS INT INC.
 XX PI Lim-wilby M, Levy OE, Brunck TK;
 XX DR WPI; 2002-361643/39.
 XX CC Novel peptide compound having hepatitis C virus protease inhibitory
 XX PT activity useful for treating disorders associated with hepatitis C
 XX PT virus protease -
 XX PS Claim 17; Page 64; 69pp; English.
 XX CC The sequence represents a peptide compound of the invention having
 XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 XX CC invention are alpha-ketoamide peptide analogues. The peptides have
 XX CC virucide activity, and are useful for treating and in the manufacture of
 XX CC a medicament to treat disorders associated with HCV protease. A
 XX CC pharmaceutical composition comprising the peptide as an active ingredient
 XX CC is useful for treating disorders associated with hepatitis C virus.
 XX SQ Sequence 11 AA;
 XX Query Match 96.4%; Score 54; DB 23; Length 11;
 XX Best Local Similarity 100.0%; Pred. No. 0.00045;
 XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 BEVVPXGMHYS 11
 DB |||||
 1 BEVVPXGMHYS 11

RESULT 2
 ABB80527
 ID ABB80527 standard; peptide; 11 AA.
 XX AC ABB80527;
 XX DT 08-OCT-2002 (first entry)
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #7.
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 XX KW virucide.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 XX 31-JAN-2002.
 XX PF 19-JUL-2001; 2001WO-US23169.
 XX PR 21-JUL-2000; 2000US-220101P.
 XX PA (CORV-) CORVAS INT INC.
 XX PI Lim-wilby M, Levy OE, Brunck TK;
 XX DR WPI; 2002-361643/39.
 XX CC Novel peptide compound having hepatitis C virus protease inhibitory

PT activity useful for treating disorders associated with hepatitis C
 PT virus protease -
 XX Claim 17; Page 64; 69pp; English.
 XX CC The sequence represents a peptide compound of the invention having
 XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 XX CC invention are alpha-ketoamide peptide analogues. The peptides have
 XX CC virucide activity, and are useful for treating and in the manufacture of
 XX CC a medicament to treat disorders associated with HCV protease. A
 XX CC pharmaceutical composition comprising the peptide as an active ingredient
 XX CC is useful for treating disorders associated with hepatitis C virus.
 XX SQ Sequence 11 AA;
 XX Query Match 96.4%; Score 54; DB 23; Length 11;
 XX Best Local Similarity 100.0%; Pred. No. 0.00045;
 XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 BEVVPXGMHYS 11
 DB |||||
 1 BEVVPXGMHYS 11

RESULT 3
 ABB80558
 ID ABB80558 standard; peptide; 11 AA.
 XX AC ABB80558;
 XX DT 08-OCT-2002 (first entry)
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #38.
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 XX KW virucide.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Modified-site 8 /note= "Oxymethionine"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 XX 31-JAN-2002.
 XX PF 19-JUL-2001; 2001WO-US23169.
 XX PR 21-JUL-2000; 2000US-220101P.
 XX PA (CORV-) CORVAS INT INC.
 XX PI Lim-wilby M, Levy OE, Brunck TK;
 XX DR WPI; 2002-361643/39.
 XX CC Novel peptide compound having hepatitis C virus protease inhibitory
 XX PT activity useful for treating disorders associated with hepatitis C
 XX PT virus protease -
 XX PS Claim 17; Page 65; 69pp; English.
 XX CC The sequence represents a peptide compound of the invention having
 XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 XX CC invention are alpha-ketoamide peptide analogues. The peptides have

CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;
Query Match 96.4%; Score 54; DB 23; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEVVPXGMHYS 11
| | | | |
Db 1 EEVVPXGMHYS 11

RESULT 4
ABB0560
ID ABB0560 standard; peptide; 11 AA.

XX AC ABB0560;

XX DT 08-OCT-2002 (first entry)

XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #40.

XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX virucide.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"

FT Misc-difference 8 /note= "D-form residue"

FT Modified-site 8 /note= "Oxymethionine"

FT Misc-difference 9 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

XX WO200208251-A2.

XX PD 31-JAN-2002.

XX PF 19-JUL-2001; 2001WO-US23169.

XX PR 21-JUL-2000; 2000US-220101P.

XX PA (CORV-) CORVAS INT INC.

XX PI Lim-wilby M, Levy OB, Brunck TK;

XX DR WPI; 2002-361643/39.

XX PT Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease -

XX PS Claim 17; Page 65; 69pp; English.

XX CC The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;
Query Match 96.4%; Score 54; DB 23; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEVVPXGMHYS 11
| | | | |
Db 1 EEVVPXGMHYS 11

RESULT 5

ABB0537

ID ABB0537 standard; peptide; 11 AA.

XX AC ABB0537;

XX DT 08-OCT-2002 (first entry)

XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #17.

XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX virucide.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"

FT Misc-difference 9 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

XX WO200208251-A2.

XX PD 31-JAN-2002.

XX PF 19-JUL-2001; 2001WO-US23169.

XX PR 21-JUL-2000; 2000US-220101P.

XX PA (CORV-) CORVAS INT INC.

XX PI Lim-wilby M, Levy OE, Brunck TK;

XX DR WPI; 2002-361643/39.

XX PT Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease -

XX PS Claim 17; Page 64; 69pp; English.

XX CC The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;
Query Match 87.5%; Score 49; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.0043;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEVVPXGMHYS 11
| | | | |


```
XX Key Location/Qualifiers
FH Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT Modified-site 6 /note= "residue 7"
FT
FT Misc-difference 9 /note= "D-form residue"
FT Modified-site 11 /note= "C-terminal amide"
FT
XX WO200208251-A2.
PN
XX 31-JAN-2002.
XX
XX 19-JUL-2001; 2001WO-US23169.
XX
XX 21-JUL-2000; 2000US-220101P.
XX (CORV-) CORVAS INT INC.
XX
XX Lim-wilby M, Levy OE, Brunck TK;
XX WPI; 2002-361643/39.
XX
XX Novel peptide compound having hepatitis C virus protease inhibitory
XX activity useful for treating disorders associated with hepatitis C
XX virus protease -
XX
XX Claim 17; Page 65; 69pp; English.
XX
XX The sequence represents a peptide compound of the invention having
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX invention are alpha-ketoamide peptide analogues. The peptides have
XX virucide activity, and are useful for treating and in the manufacture of
XX a medicament to treat disorders associated with HCV protease. A
XX pharmaceutical composition comprising the peptide as an active ingredient
XX is useful for treating disorders associated with hepatitis C virus.
XX
XX Sequence 11 AA;
SQ
Query Match 85.7%; Score 48; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. NO. 0.0067;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPXGMHYS 11
DB ||||| |||
1 EEVVPXGSHYS 11
RESULT 9
ABB80554
ID ABB80554 standard; peptide; 11 AA.
XX
XX ABB80534;
AC
XX 08-OCT-2002 (first entry)
DT
XX
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #34.
DE
XX
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX
XX Synthetic.
OS
XX Key Location/Qualifiers
FH Modified-site 1 /note= "N-terminal acetyl"
FT
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"
FT
FT Misc-difference 8
PN WO200208251-A2.
XX
```

```
FT Modified-site 11 /note= "D-form residue"
FT FT /note= "C-terminal amide"
XX
XX WO200208251-A2.
PN
XX 31-JAN-2002.
XX
XX 19-JUL-2001; 2001WO-US23169.
XX
XX 21-JUL-2000; 2000US-220101P.
XX (CORV-) CORVAS INT INC.
XX
XX Lim-wilby M, Levy OE, Brunck TK;
XX WPI; 2002-361643/39.
XX
XX Novel peptide compound having hepatitis C virus protease inhibitory
XX activity useful for treating disorders associated with hepatitis C
XX virus protease -
XX
XX Claim 17; Page 65; 69pp; English.
XX
XX The sequence represents a peptide compound of the invention having
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX invention are alpha-ketoamide peptide analogues. The peptides have
XX virucide activity, and are useful for treating and in the manufacture of
XX a medicament to treat disorders associated with HCV protease. A
XX pharmaceutical composition comprising the peptide as an active ingredient
XX is useful for treating disorders associated with hepatitis C virus.
XX
XX Sequence 11 AA;
SQ
Query Match 85.7%; Score 48; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. NO. 0.0067;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPXGMHYS 11
DB ||||| |||
1 EEVVPXGSHYS 11
RESULT 10
ABB80555
ID ABB80555 standard; peptide; 11 AA.
XX
XX ABB80555;
AC
XX 08-OCT-2002 (first entry)
DT
XX
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #35.
DE
XX
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX
XX Synthetic.
OS
XX Key Location/Qualifiers
FH Modified-site 1 /note= "N-terminal acetyl"
FT
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"
FT
FT Misc-difference 8 /note= "D-form residue"
FT
FT Misc-difference 9 /note= "D-form residue"
FT
FT Modified-site 11 /note= "C-terminal amide"
XX
XX WO200208251-A2.
XX
```

PD 31-JAN-2002.
 XX
 PF 19-JUL-2001; 2001WO-US23169.
 XX
 PR 21-JUL-2000; 2000US-220101P.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX
 XX WPI; 2002-361643/39.
 XX
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease -
 XX
 XX Claim 17; Page 65; 69pp; English.
 XX
 CC The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 XX SQ Sequence 11 AA;
 XX
 Query Match 85.7%; Score 48; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.0067;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGMHYS 11
 DB 1 EEVVPXGSHYS 11
 XX
 RESULT 11
 ABB80531
 ID ABB80531 standard; peptide; 11 AA.
 XX
 AC ABB80531;
 XX
 DT 08-OCT-2002 (first entry)
 XX
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #11.
 XX
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Modified-site 11 /note= "C-terminal amide"
 XX
 XX WO200208251-A2.
 XX
 XX 31-JAN-2002.
 XX
 XX 19-JUL-2001; 2001WO-US23169.
 XX
 XX 21-JUL-2000; 2000US-220101P.
 XX
 XX (CORV-) CORVAS INT INC.
 XX
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX
 XX WPI; 2002-361643/39.
 XX
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease -
 XX
 XX Claim 17; Page 65; 69pp; English.

DR WPI; 2002-361643/39.
 XX
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease -
 XX
 XX Claim 17; Page 64; 69pp; English.
 XX
 CC The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 XX SQ Sequence 11 AA;
 XX
 Query Match 82.1%; Score 46; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.017;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGMHYS 11
 DB 1 EEVVPXGSHYS 11
 XX
 RESULT 12
 ABB80532
 ID ABB80532 standard; peptide; 11 AA.
 XX
 AC ABB80532;
 XX
 DT 08-OCT-2002 (first entry)
 XX
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #12.
 XX
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 9 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 XX
 XX WO200208251-A2.
 XX
 XX 31-JAN-2002.
 XX
 XX 19-JUL-2001; 2001WO-US23169.
 XX
 XX 21-JUL-2000; 2000US-220101P.
 XX
 XX (CORV-) CORVAS INT INC.
 XX
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX
 XX WPI; 2002-361643/39.
 XX
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease -
 XX
 XX Claim 17; Page 64; 69pp; English.

CC The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.

SQ Sequence 11 AA;

Query Match 82.1%; Score 46; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.017; 1; Indels 0; Gaps 0;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11

DB 1 EEVVPXGMHYS 11

RESULT 13

ABB80521
 ID ABB80521 standard; peptide; 11 AA.

XX AC ABB80521;

DT 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #1.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 virucide.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with

FT Modified-site 11 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease

XX Claim 17; Page 64; 69pp; English.

CC The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 80.4%; Score 45; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.026;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11

DB 1 EEVVPXGMHYS 11

RESULT 14

ABB80522
 ID ABB80522 standard; peptide; 11 AA.

XX AC ABB80522;

DT 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #2.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 virucide.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with

FT Misc-difference 9 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease

XX Claim 17; Page 64; 69pp; English.

CC The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.

SQ Sequence 11 AA;

Query Match 80.4%; Score 45; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.026;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11

DB 1 EEVVPXGMHYS 11

RESULT 15
 ABB80524
 ID ABB80524 standard; peptide; 11 AA.
 XX
 AC ABB80524;
 XX
 DT 08-OCT-2002 (first entry)
 XX
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #4.
 XX
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 virucide.
 XX
 OS Synthetic.
 XX
 PH Key Location/Qualifiers
 FT Modified-site 1
 FT FT /note= "N-terminal acetyl"
 FT Modified-site 6
 FT FT /note= "Norvalyl carbonyl forming keto-amide linkage with
 residue 7"
 FT Misc-difference 9
 FT FT /note= "D-form residue"
 FT Modified-site 11
 FT FT /note= "C-terminal amide"
 XX
 PN WO200208251-A2.
 XX
 PD 31-JAN-2002.
 XX
 PF 19-JUL-2001; 2001WO-US23169.
 XX
 PR 21-JUL-2000; 2000US-220101P.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 PI Lim-wilby M, Levy OE, Brunck TK;
 XX
 DR WPI; 2002-361643/39.
 XX
 PT Novel peptide compound having hepatitis C virus protease inhibitory
 activity useful for treating disorders associated with hepatitis C
 virus protease -
 XX
 PS Claim 17; Page 64; 69pp; English.
 XX
 CC The sequence represents a peptide compound of the invention having
 hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 invention are alpha-ketoamide peptide analogues. The peptides have
 virucide activity, and are useful for treating and in the manufacture of
 a medicament to treat disorders associated with HCV protease. A
 pharmaceutical composition comprising the peptide as an active ingredient
 is useful for treating disorders associated with hepatitis C virus.
 XX
 SQ Sequence 11 AA;
 Query Match 80.4%; Score 45; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.026;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPXGMHYS 11
 |||||
 Db 1 EEVVPXGMDYS 11

Search completed: June 4, 2003, 13:11:13
 Job time : 29.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:07:09 ; Search time 22.5 Seconds
(without alignments)
100.734 Million cell updates/sec

Title: AUDET-909-2

Perfect score: 56

Sequence: 1 eavpvgmhys 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_xvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	67.9	1063	Q8RG86	Q8rg86 fusobacteri
2	37	66.1	308	Q9X2E2	Q9x2e2 thermotoga
3	37	66.1	322	17 Q9HLH8	Q9hlh8 thermoplas
4	37	66.1	1057	16 Q99UR5	Q99ur5 staphylococ
5	36	64.3	208	2 Q46486	Q46486 corynebacte
6	36	64.3	252	16 Q8VWPL	Q8vwp1 anabaena sp
7	36	64.3	819	10 Q9AVK4	Q9avk4 pisum sativ
8	35	62.5	139	2 Q57489	Q57489 bacteroides
9	35	62.5	156	16 Q9PC35	Q9pc35 xylella fas
10	35	62.5	233	10 Q40479	Q40479 nicotiana t
11	35	62.5	237	10 Q9LW50	Q9lw50 nicotiana s
12	35	62.5	317	9 Q38317	Q38317 lactobacill
13	35	62.5	425	5 Q9XVK4	Q9xvk4 caenorhabdi
14	35	62.5	510	10 Q9SA71	Q9sa71 arabidopsis
15	35	62.5	2042	17 Q8TZ07	Q8tz07 methanopyru
16	34	60.7	264	17 Q27902	Q27902 methanobact

17	34	60.7	279	16	Q9RXN9	Q9rxn9 deinococcus
18	34	60.7	285	16	Q98HU6	Q98hu6 rhizobium l
19	34	60.7	350	16	Q9RW92	Q9rw92 deinococcus
20	34	60.7	355	16	Q9XAM3	Q9xam3 streptomyce
21	34	60.7	360	17	Q27679	Q27679 methanobact
22	34	60.7	425	16	Q9KBA1	Q9kba1 bacillus na
23	34	60.7	495	5	O16912	O16912 caenorhabdi
24	34	60.7	637	5	Q9YQY6	Q9yqy6 drosophila
25	34	60.7	678	12	Q9E1X6	Q9elx6 cercopithe
26	34	60.7	748	12	Q9YR01	Q9yr01 ranid herpe
27	34	60.7	1028	16	Q8YJ11	Q8yj11 brucella me
28	34	60.7	2438	5	Q9VQL7	Q9vql7 drosophila
29	33	58.9	156	3	Q12479	Q12479 saccharomyc
30	33	58.9	216	16	Q9RDC1	Q9rdc1 streptomyce
31	33	58.9	257	8	O99011	O99011 prototheca
32	33	58.9	273	4	Q96N44	Q96n44 homo sapien
33	33	58.9	306	16	Q8XVQ7	Q8xvq7 ralstonia s
34	33	58.9	367	16	Q916J1	Q916j1 pseudomonas
35	33	58.9	387	16	Q98FX1	Q98fx1 rhizobium l
36	33	58.9	426	3	Q9HG99	Q9hg99 mortierella
37	33	58.9	441	16	Q9KTH7	Q9kth7 vibrio chol
38	33	58.9	466	16	Q9PKH7	Q9pkh7 chlamydia m
39	33	58.9	466	16	O84218	O84218 chlamydia t
40	33	58.9	534	17	O29966	O29966 archaeoglob
41	33	58.9	542	12	Q9WCW0	Q9wcw0 avian infec
42	33	58.9	545	10	Q9T025	Q9t025 arabidopsis
43	33	58.9	556	12	Q9QGT4	Q9qgt4 avian infec
44	33	58.9	558	12	Q9QGT3	Q9qgt3 avian infec
45	33	58.9	584	12	Q9QGT1	Q9qgt1 avian infec

ALIGNMENTS

RESULT 1

Q8RG86 ID Q8RG86 PRELIMINARY; PRT; 1063 AA.
AC Q8RG86;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5).
GN FN0422.
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11889109;
RA Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Lasieva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA Forstein M., Kyrpides N., Overbeek R.,
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586";
RL J. Bacteriol. 184:2005-2018(2002).
DR EMBL; AE010554; AAL94625.1; -.
KW Ligase; Complete proteome.
SQ SEQUENCE 1063 AA; 118008 MW; 39700E10B7CCE411 CRC64;

Query Match 67.9%; Score 38; DB 16; Length 1063;
Best Local Similarity 60.0%; Pred. No. 48;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EVVPXGMHYS 11

Db 195 EIVPGLNYS 204

RESULT 2

Q9X2E2

```

ID Q9X2E2 PRELIMINARY; PRT; 308 AA.
AC Q9X2E2;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE FISH protease activity modulator HFLK.
GN TM1822.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Uterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RL genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
DR EMBL; AE001819; AAD36885.1; -.
DR TIGR; TM1822; -.
DR InterPro; IPR001107; Band 7.
DR InterPro; IPR003130; GED_7.
DR InterPro; IPR001972; Stomatatin.
DR Pfam; PF01145; Band_7; 1.
DR PRINTS; PR00721; STOMATIN.
DR SMART; SM00302; GED; 1.
DR SMART; SM00244; PHB; 1.
KW Protease; Complete proteome.
SQ SEQUENCE 308 AA; 34778 MW; ADE03603E5101A9D CRC64;

Query Match 66.1%; Score 37; DB 16; Length 308;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMHY 10
DB 41 VVPSGIHY 48

RESULT 3
Q9HLH8 PRELIMINARY; PRT; 322 AA.
ID Q9HLH8
AC Q9HLH8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Glucose-fructose oxidoreductase related protein.
GN TA0250.
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT acidophilum.";
RL Nature 407:508-513 (2000).
DR EMBL; AL445063; CAC11395.1; -.
DR InterPro; IPR000683; GFO_IDH_MCA.
DR Pfam; PF01408; GFO_IDH_MCA; 1.
KW Complete proteome.
SQ SEQUENCE 322 AA; 36918 MW; B8C239E71009D167 CRC64;

Query Match 66.1%; Score 37; DB 17; Length 322;

```

```

Best Local Similarity 75.0%; Pred. No. 22;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMHY 10
DB 66 VVPDGLHY 73

RESULT 4
Q99UR5 PRELIMINARY; PRT; 1057 AA.
ID Q99UR5
AC Q99UR5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Carbamoyl-phosphate synthase large chain.
GN PRAB OR SAV1203 OR SA1046.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OC Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Murakami A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003361; BAB57365.1; -.
DR EMBL; AP003132; BAB42298.1; -.
DR HSP; P00968; ICSO.
DR InterPro; IPR000901; CPSase.
DR InterPro; IPR004362; MGS_like.
DR InterPro; IPR000189; SHprot_acsite.
DR Pfam; PF00289; CPSase_L_chain; 2.
DR Pfam; PF02786; CPSase_L_D2; 2.
DR Pfam; PF02787; CPSase_L_D3; 1.
DR Pfam; PF02142; MGS; 1.
DR PRINTS; PR00098; CPSASE.
DR PROSITE; PS00866; CPSASE_1; 2.
DR PROSITE; PS00867; CPSASE_2; UNKNOWN_2.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 1057 AA; 117171 MW; E3E179EF0591F0F8 CRC64;

Query Match 66.1%; Score 37; DB 16; Length 1057;
Best Local Similarity 60.0%; Pred. No. 77;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMHY 11
DB 190 EIVSNGLHY 199

RESULT 5
Q46486 PRELIMINARY; PRT; 208 AA.
ID Q46486
AC Q46486;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Hypothetical 23.0 kDa protein (GcrA).
GN GCR.
OS Corynebacterium xerosis, and
OS Corynebacterium striatum.

```

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Corynebacteriaceae;
 OC Corynebacterium.
 OX NCBI_TaxID=1725, 43770;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=C.xerosus; STRAIN=M82B;
 RX MEDLINE=96117603; PubMed=8559800;
 RA Tauch A., Kassing F., Kalinowski J., Puhler A.;
 RT "The Corynebacterium xerosus composite transposon Tn5432 consists of
 RT two identical insertion sequences, designated IS249, flanking the
 RT erythromycin resistance gene ermC.";
 RL Plasmid 34:119-131(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=C.striatum; STRAIN=M82B;
 RX MEDLINE=20194806; PubMed=10732668;
 RA Tauch A., Krieff S., Kalinowski J., Puhler A.;
 RT "The 51,409-bp R-plasmid pTP10 from the multiresistant clinical
 RT isolate Corynebacterium striatum M82B is composed of DNA segments
 RT initially identified in soil bacteria and in plant, animal, and human
 RT pathogens.";
 RL Mol. Gen. Genet. 263:1-11(2000).
 DR EMBL; U21300; AAC95478.1; -;
 DR EMBL; AF024666; AAG03390.1; -;
 KW Hypothetical protein; Plasmid.
 SQ SEQUENCE 208 AA; 23012 MW; F1504BE1ECD885A6 CRC64;

Query Match 64.3%; Score 36; DB 2; Length 208;
 Best Local Similarity 50.0%; Pred. No. 22;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMHVS 11
 :||:|:|:
 Db 130 DVIPEKHYA 139

RESULT 6
 Q8YWP1 PRELIMINARY; PRT; 252 AA.
 ID Q8YWP1
 AC Q8YWP1;
 DT 01-MAR-2002 (TReMBLrel. 20, Created)
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Hypothetical protein Alr1563.
 GN ALR1563.
 OS Anabaena sp. (strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 OX NCBI_TaxID=103690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21595285; PubMed=11759840;
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
 RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 RT cyanobacterium Anabaena sp. strain PCC 7120.";
 RL DNA Res. 8:205-213(2001).
 DR EMBL; AF003586; BAB77929.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 252 AA; 28831 MW; 925572DA5D1CA519 CRC64;

Query Match 64.3%; Score 36; DB 16; Length 252;
 Best Local Similarity 50.0%; Pred. No. 27;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVVPXGMHVS 10
 :||:|:|:
 Db 235 EMIVPAGLHF 244

RESULT 7
 Q9AVK4 PRELIMINARY; PRT; 819 AA.
 ID Q9AVK4
 AC Q9AVK4;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE SCARECROW.
 GN PSSCR.
 OS Pisum sativum (Garden pea).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
 OX NCBI_TaxID=3888;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. ALASKA;
 RX MEDLINE=21231727; PubMed=11333309;
 RA Sassa N., Matsushita Y., Nakamura T., Nyunoya H.;
 RT "The Molecular Characterization and in situ Expression Pattern of Pea
 RT SCARECROW Gene.";
 RL Plant Cell Physiol. 42:385-394(2001).
 DR EMBL; AB048713; BAB39155.1; -;
 DR InterPro; IPR001444; Flag_bb_rod.
 DR InterPro; IPR005202; GRAS.
 DR Pfam; PF03514; GRAS; 1.
 DR PROSITE; PS00588; FLAGELLIN_BB_ROD; UNKNOWN 1.
 SQ SEQUENCE 819 AA; 90372 MW; 41B67BD6DC72ADFA CRC64;

Query Match 64.3%; Score 36; DB 10; Length 819;
 Best Local Similarity 45.5%; Pred. No. 94;
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 BEVVPXGMHVS 11
 :||:|:|:
 Db 343 DDVVPISLHFS 353

RESULT 8
 Q57489 PRELIMINARY; PRT; 139 AA.
 ID Q57489
 AC Q57489;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE DNA ligase (Fragment).
 OS Bacteroides nodosus (Dichelobacter nodosus).
 OC Bacteria; Proteobacteria; gamma subdivision; Cardiobacteriaceae;
 OC Dichelobacter.
 OX NCBI_TaxID=870;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96020672; PubMed=7476204;
 RA Moses E.K., Good R.T., Sinistaj M., Billington S.J., Langford C.J.,
 RA Rood J.I.;
 RT "A multiple site-specific DNA-inversion model for the control of Omp1
 RT phase and antigenic variation in Dichelobacter nodosus.";
 RL Mol. Microbiol. 17:183-196(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96257263; PubMed=8654969;
 RA Billington S.J., Sinistaj M., Cheetham B.F., Ayres A., Moses E.K.,
 RA Katz M.E., Rood J.I.;
 RT "Identification of a native Dichelobacter nodosus plasmid and
 RT implications for the evolution of the vap regions.";
 RL Gene 172:111-116(1996).
 DR EMBL; U02462; AAB12366.1; -;
 DR InterPro; IPR001357; BRCT.
 DR Pfam; PF00533; BRCT; 1.
 DR SMART; SM00292; BRCT; 1.
 DR PROSITE; PS0172; BRCT; 1.
 KW Ligase.

FT NON_TER 1 1

SQ SEQUENCE 139 AA; 15052 MW; E0E110AA4B7D4708 CRC64;
 Query Match 62.5%; Score 35; DB 2; Length 139;
 Best Local Similarity 55.6%; Pred. No. 24;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 Qy 3 VVPXGMHYS 11
 Db 21 IVPAVGWHS 29
 RESULT 9
 Q9PC35 PRELIMINARY; PRT; 156 AA.
 AC Q9PC35
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hypothetical protein Xf1950.
 GN Xf1950.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xylella.
 OX NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9A5C;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Bonaccorsi A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa";
 RL Nature 406:151-159(2000).
 DR EMBL; AB004014; AAF84752.1; --
 DR InterPro; IPR002545; Chw.
 DR SMART; SM00260; Chw; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 156 AA; 17144 MW; D8358619C6671A5D CRC64;
 Query Match 62.5%; Score 35; DB 16; Length 156;
 Best Local Similarity 55.6%; Pred. No. 27;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 EEVVPXGMH 9
 Db 119 EEILPGVH 127
 RESULT 10
 Q40479 PRELIMINARY; PRT; 233 AA.
 AC Q40479;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE EREBP-2.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BY4; TISSUE=LEAF;
 RX MEDLINE=95276459; PubMed=7756828;
 RA Ohme-Takagi M., Shinshi H.;
 RT "Ethylene-inducible DNA binding proteins that interact with an
 RT ethylene-responsive element";
 RL Plant Cell 7:173-182(1995).
 DR EMBL; D38126; BAA07324.1; --
 DR HSSP; O80337; 2GCC.
 DR TRANSFAC; T02654; --
 DR InterPro; IPR001471; TF ERF.
 DR Pfam; PF00847; AP2-domain; 1.
 DR PRINTS; PR00367; ETHRSPLEMT.
 DR ProDom; PD001423; TF AP2; 1.
 DR SMART; SM00380; AP2; 1.
 SQ SEQUENCE 233 AA; 25563 MW; 6CD16783582C0CB5 CRC64;
 Query Match 62.5%; Score 35; DB 10; Length 233;
 Best Local Similarity 60.0%; Pred. No. 41;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 EEVVPXGMHY 10
 Db 90 QAVVPKGRHY 99
 RESULT 11
 Q9LW50 PRELIMINARY; PRT; 237 AA.
 AC Q9LW50;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Ethylene-responsive element binding factor.
 GN NSERF2.
 OS Nicotiana sylvestris (Wood tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4096;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20399450; PubMed=10945353;
 RA Kitajima S., Koyama T., Ohme-Takagi M., Shinshi H., Sato F.;
 RT "Characterization of gene expression of NSERFs, transcription factors
 RT of basic PR genes from Nicotiana sylvestris";
 RL Plant Cell Physiol. 41:817-824(2000).
 DR EMBL; AB016264; BAA97122.1; --
 DR HSSP; O80337; 2GCC.
 DR InterPro; IPR001471; TF ERF.
 DR Pfam; PF00847; AP2-domain; 1.
 DR PRINTS; PR00367; ETHRSPLEMT.
 DR ProDom; PD001423; TF AP2; 1.
 DR SMART; SM00380; AP2; 1.
 SQ SEQUENCE 237 AA; 26243 MW; 01BC3EBE51E46298 CRC64;
 Query Match 62.5%; Score 35; DB 10; Length 237;
 Best Local Similarity 60.0%; Pred. No. 41;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 EEVVPXGMHY 10

Db 94 QAVVPGRRHY 103
: ||| |||

RESULT 12

Q38317 ID Q38317 PRELIMINARY; PRT; 317 AA.
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Lysin.
GN Lys.
OS Lactobacillus bacteriophage phi adh.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=12417;
RN (1)
RP SEQUENCE FROM N.A.
RA Altermann E.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99384014; PubMed=10452953;
RA Altermann E., Klein J., Henrich B.;
RT "Primary structure and features of the genome of the Lactobacillus gasseri temperate bacteriophage phi-adh.";
RL Gene 236:333-346(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=95138034; PubMed=7836307;
RA Henrich B., Binihofer B., Blaesi U.;
RT "Primary structure and functional analysis of the lysis genes of Lactobacillus gasseri bacteriophage phi-adh.";
RL J. Bacteriol. 177:723-732(1995).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=93231538; PubMed=8472961;
RA Frenaux C., De Antoni G., Raya R., Klaenhammer T.;
RT "Genetic organization and sequence of the region encoding integrative functions from Lactobacillus gasseri temperate bacteriophage phi-adh.";
RL Gene 126:61-66(1993).
RN [5]
RP SEQUENCE FROM N.A.
RA Engel G., Altermann B., Klein J., Henrich B.;
RT "Structure of a genome region of the Lactobacillus gasseri temperate phage phi adh covering a repressor gene and cognate promoters.";
RL Gene 210:67-70(1998).
DR EMBL; AJ131519; CAB52540.1; -.
DR InterPro; IPR002053; GH_25.
DR Pfam; PF01183; Glyco_hydro_25; 1.
DR ProDom; PD004820; GH_25; 1.
DR SMART; SM00287; SH3b; 1.
SQ SEQUENCE 317 AA; 34703 MW; 9FF2715EE43561C7 CRC64;

Query Match 62.5%; Score 35; DB 9; Length 317;
Best Local Similarity 66.7%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMHYS 11
: ||| |||

RESULT 13

Q9XVK4 ID Q9XVK4 PRELIMINARY; PRT; 425 AA.
AC Q9XVK4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE R10D12.10 protein.

Query Match 62.5%; Score 35; DB 9; Length 317;
Best Local Similarity 66.7%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMHYS 11
: ||| |||

GN R10D12.10.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Percy C.M.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z81109; CAB03241.1; -.
DR InterPro; IPR000719; Euk_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Transferase.
SQ SEQUENCE 425 AA; 49410 MW; 5D96E29B08C8E9D6 CRC64;

Query Match 62.5%; Score 35; DB 5; Length 425;
Best Local Similarity 50.0%; Pred. No. 76;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10
: ||| |||

RESULT 14

Q9SA71 ID Q9SA71 PRELIMINARY; PRT; 510 AA.
AC Q9SA71;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE T518.1 protein.
GN T518.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,
Li J., Kremenetskaia I., Luros J., Ngan I., Gonzalez A., Altafi H.,
Araujo R., Chao Q., Conn L., Conway A.B., Dunn P., Hansen N.,
Huizar L., Kim C., Palm C., Rowley D., Shinn P., Walker M.,
Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;
RT "Arabidopsis thaliana chromosome 1 BAC T518 sequence.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; AC007060; AAD25743.1; -.
KW Transmembrane.
SQ SEQUENCE 510 AA; 55089 MW; 0DC63CC41F4712F8 CRC64;

Query Match 62.5%; Score 35; DB 10; Length 510;
Best Local Similarity 60.0%; Pred. No. 92;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10
: ||| |||

RESULT 15

Q8T207

ID Q8TZ07 PRELIMINARY; PRT; 2042 AA.
 AC Q8TZ07;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Predicted protein of the CoN/MG-chelataase family.
 GN MK0134.
 OS Methanopyrus kandleri.
 OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
 OC Methanopyrus.
 OX NCBI_TaxID=2320;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AV19 / DSM 6324 / JCM 9639;
 RX MEDLINE=21927647; PubMed=11930014;
 RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,
 RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
 RA Natile D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
 RA Malykh A.G., Koonin E.V., Kozyavkin S.A.;
 RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
 and monophyly of archaeal methanogens.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649 (2002).
 DR EMBL; AE010313; AAM01351.1; -.
 KW Complete proteome.
 SQ SEQUENCE 2042 AA; 227716 MW; 83B6E092A62C112A CRC64;

Query Match 62.5%; Score 35; DB 17; Length 2042;
 Best Local Similarity 75.0%; Pred. No. 4e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 EVVPXGMH 9
 Db 924 EVVPIGLH 931
 |||||
 |||||

Search completed: June 4, 2003, 13:13:25
 Job time : 24.5 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:04:34 ; Search time 6.25 Seconds
(without alignments)
72.998 Million cell updates/sec

Title: AUDET-909-2

Perfect score: 56

Sequence: 1 eevvpxgmhys 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.4

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	38	67.9	1058	1 CARB_FUSNN	Q8rg86 fusobacteri
2	37	66.1	1057	1 CARB_STAAM	Q99ur5 staphylococ
3	37	66.1	1057	1 CARB_STAAM	P58940 staphylococ
4	36	64.3	102	1 Y1LK_TYDVA	P31619 tobacco yel
5	36	64.3	460	1 UMEI_YEAST	Q03010 saccharomyc
6	36	64.3	743	1 YK47_YEAST	P36148 saccharomyc
7	35	62.5	227	1 ID11_MESAU	O35586 mesocricetu
8	35	62.5	308	1 GABR_WETTH	O28806 methanobact
9	34	60.7	426	1 SLSI_YARLI	Q99158 yarrowia li
10	33	58.9	513	1 PHSL_DESBA	P13065 desulfovibr
11	33	58.9	627	1 MUTL_BACSU	P49850 bacillus su
12	33	58.9	1188	1 KPBA_CAEEL	P34335 caenorhabdi
13	33	58.9	1396	1 VLF_BPT5	P13390 bacterioph
14	32	57.1	126	1 TKNK_BOVIN	P08858 bos taurus
15	32	57.1	267	1 RR2_CHLVU	P56351 chlorella v
16	32	57.1	288	1 CGD2_RAT	Q04827 rattus norv
17	32	57.1	289	1 CGD2_HUMAN	P30279 homo sapien
18	32	57.1	289	1 CGD2_MOUSE	P30280 mus musculu
19	32	57.1	291	1 CGD1_BRARE	Q90459 brachydanio
20	32	57.1	291	1 CGD1_XENLA	P50755 xenopus lae
21	32	57.1	291	1 CGD2_CHICK	P49706 gallus gall
22	32	57.1	291	1 CGD2_CHICK	P53782 xenopus lae
23	32	57.1	292	1 CGD1_CHICK	P55169 gallus gall
24	32	57.1	292	1 CGD3_HUMAN	P30281 homo sapien
25	32	57.1	295	1 CGD1_HUMAN	P24385 homo sapien
26	32	57.1	295	1 CGD1_MOUSE	P25322 mus musculu
27	32	57.1	295	1 CGD1_RAT	P39948 rattus norv
28	32	57.1	341	1 HYPE_AZOVI	P40595 azotobacter
29	32	57.1	353	1 T2BA_BACAR	P19887 bacillus an
30	32	57.1	759	1 SCT1_YEAST	P32784 saccharomyc
31	32	57.1	877	1 SULH_SCHPO	O74377 schizosacch
32	32	57.1	1401	1 RPOC_VIBCH	Q9kv29 vibrio chol
33	32	57.1	1717	1 ZEP1_HUMAN	P15822 homo sapien

RESULT 1

ID	CARB_FUSNN	STANDARD;	PRT;	1058 AA.
AC	Q8RG86;			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-phosphate synthetase ammonia chain).			
DE	CARB OR FN0422.			
GN	Fusobacterium nucleatum (subsp. nucleatum).			
OS	Bacteria; Fusobacteria; Fusobacterium.			
OX	NCBI_TaxID=76856;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 25586;			
RC	MEDLINE=21886394; PubMed=11889109;			
RA	Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A., Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L., Vasieva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A., Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R., Fonstein N., Kyripides N., Overbeek R.;			
RA	"Genome sequence and analysis of the oral bacterium Fusobacterium nucleatum strain ATCC 25586";			
RT	J. Bacteriol. 184:2003-2018(2002).			
RL	-!- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP + phosphate + L-glutamate + carbamoyl phosphate.			
CC	-!- COFACTOR: Binds three manganese ions (By similarity).			
CC	-!- PATHWAY: Arginine biosynthesis.			
CC	-!- PATHWAY: Pyrimidine biosynthesis; first step.			
CC	-!- SUBUNIT: Composed of two chains; the small (or glutamine) chain promotes the hydrolysis of glutamine to ammonia, which is used by the large (or ammonia) chain to synthesize carbamoyl phosphate (By similarity).			
CC	-!- SIMILARITY: BELONGS TO THE CARB FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; AE010554; AAL94625.1; ALT_INIT.			
DR	InterPro; IPR005483; CPase_L.			
DR	InterPro; IPR005479; CPase_L_D2.			
DR	InterPro; IPR005480; CPase_L_D3.			
DR	InterPro; IPR005481; CPase_L_N.			
DR	InterPro; IPR004362; MGS like.			
DR	Pfam; PF00289; CPase_L chain; 2.			
DR	Pfam; PF02786; CPase_L_D2; 2.			
DR	Pfam; PF02787; CPase_L_D3; 1.			
DR	Pfam; PF02142; MGS; 1.			
DR	PRINTS; PR00098; CPASE.			
DR	PROSITE; PS00866; CPASE_1; 2.			

P20273 homo sapien
P17280 chimpanzee
P97885 rattus norv
O58349 methanococc
P10941 cryphonectr
O28354 archaeoglob
P52210 ovis aries
P44640 haemophilus
O75355 homo sapien
P40739 bacillus su
P27989 moorella th
O49230 brassica ol

ALIGNMENTS

DR PROSITE; PS00867; CPASE 2; 2.
 KW Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;
 KW ATP-binding; Manganese; Complete proteome.
 FT DOMAIN 1 401 CARBOXYPHOSPHATE SYNTHETIC DOMAIN.
 FT DOMAIN 402 546 OLIGOMERIZATION DOMAIN.
 FT DOMAIN 547 929 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.
 FT REPEAT 1 546 ALLOSTERIC DOMAIN.
 FT REPEAT 547 1058
 FT NP_BIND 153 210 ATP (POTENTIAL).
 FT NP_BIND 302 352 ATP (POTENTIAL).
 FT METAL 284 284 MANGANESE 1 (BY SIMILARITY).
 FT METAL 298 298 MANGANESE 1 AND 2 (BY SIMILARITY).
 FT METAL 300 300 MANGANESE 2 (BY SIMILARITY).
 FT METAL 820 820 MANGANESE 3 (BY SIMILARITY).
 FT METAL 832 832 MANGANESE 3 (BY SIMILARITY).
 SQ SEQUENCE 1058 AA; 117451 MW; ED7037AF77C1E39F CRC64;

Query Match 67.9%; Score 38; DB 1; Length 1058;
 Best Local Similarity 60.0%; Pred. No. 10;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 EVVPGMHYS 11
 |:| |:|
 DB 190 EIVPGLNYS 199

RESULT 2

CARB-STAAW STANDARD; PRT; 1057 AA.
 ID CARB-STAAW STANDARD; PRT; 1057 AA.
 AC Q99UR5;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-phosphate synthetase ammonia chain).
 DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-phosphate synthetase ammonia chain).
 GN CARB OR PYRAB OR SAVI203 OR SA1046.
 OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
 OS Staphylococcus aureus (strain N315).
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=158878, 158879;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mu50 / ATCC 700699, and N315;
 RX MEDLINE=21311952; PubMed=1418146;
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I., Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T., Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C., Sekinizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
 RT "whole genome sequencing of methicillin-resistant Staphylococcus aureus.";
 RL Lancet 357:1225-1240(2001).
 CC -1- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP + phosphate + L-glutamate + carbamoyl phosphate.
 CC -1- COFACTOR: Binds three manganese ions (By similarity).
 CC -1- PATHWAY: Arginine biosynthesis.
 CC -1- PATHWAY: Pyrimidine biosynthesis; first step.
 CC -1- SUBUNIT: Composed of two chains; the small (or glutamine) chain promotes the hydrolysis of glutamine to ammonia, which is used by the large (or ammonia) chain to synthesize carbamoyl phosphate (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE CARB FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AP003361; BAB57365.1; -
 DR EMBL; AP003132; BAB42298.1; -
 DR HSP; P00968; 1CS0.
 DR InterPro; IPR005483; CPase_L.
 DR InterPro; IPR005479; CPase_L_D2.
 DR InterPro; IPR005480; CPase_L_D3.
 DR InterPro; IPR005481; CPase_L_N.
 DR InterPro; IPR004362; MGS like.
 DR Pfam; PF00289; CPase_L_chain; 2.
 DR Pfam; PF02786; CPase_L_D2; 2.
 DR Pfam; PF02787; CPase_L_D3; 1.
 DR Pfam; PF02142; MGS; 1.
 DR PRINTS; PR00098; CPASE.
 DR PROSITE; PS00866; CPASE_1; 2.
 DR PROSITE; PS00867; CPASE_2; 2.
 KW Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;
 KW ATP-binding; Manganese; Complete proteome.
 FT DOMAIN 1 401 CARBOXYPHOSPHATE SYNTHETIC DOMAIN.
 FT DOMAIN 402 546 OLIGOMERIZATION DOMAIN.
 FT DOMAIN 547 929 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.
 FT DOMAIN 930 1057 ALLOSTERIC DOMAIN.
 FT REPEAT 1 546
 FT REPEAT 547 1057
 FT NP_BIND 153 210 ATP (POTENTIAL).
 FT NP_BIND 302 352 ATP (POTENTIAL).
 FT METAL 284 284 MANGANESE 1 (BY SIMILARITY).
 FT METAL 298 298 MANGANESE 1 AND 2 (BY SIMILARITY).
 FT METAL 300 300 MANGANESE 2 (BY SIMILARITY).
 FT METAL 820 820 MANGANESE 3 (BY SIMILARITY).
 FT METAL 832 832 MANGANESE 3 (BY SIMILARITY).
 SQ SEQUENCE 1057 AA; 117171 MW; E3E179EF0591F08 CRC64;
 Query Match 66.1%; Score 37; DB 1; Length 1057;
 Best Local Similarity 60.0%; Pred. No. 16;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 OY 2 EVVPGMHYS 11
 |:| |:|
 DB 190 EIVSNGHYS 199
 RESULT 3
 CARB-STAAW STANDARD; PRT; 1057 AA.
 ID CARB-STAAW STANDARD; PRT; 1057 AA.
 AC P58940;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-phosphate synthetase ammonia chain).
 DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-phosphate synthetase ammonia chain).
 GN CARB OR PYRAB OR MW1086.
 OS Staphylococcus aureus (strain MW2).
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=196620;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=22040717; PubMed=12044378;
 RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A., Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L., Yamamoto K., Hiramatsu K.;
 RT "Genome and virulence determinants of high virulence community-acquired MRSA.";
 RL Lancet 359:1819-1827(2002).
 CC -1- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP + phosphate + L-glutamate + carbamoyl phosphate.
 CC -1- COFACTOR: Binds three manganese ions (By similarity).
 CC -1- PATHWAY: Arginine biosynthesis.
 CC -1- SUBUNIT: Pyrimidine biosynthesis; first step.
 CC -1- SUBUNIT: Composed of two chains; the small (or glutamine) chain promotes the hydrolysis of glutamine to ammonia, which is used by the large (or ammonia) chain to synthesize carbamoyl phosphate (By similarity).

DR PROSITE; PS50294; WD REPEATS REGION; FALSE NEG.
 KW Transcription regulation; Meiosis; Repeat; WD repeat.
 FT REPEAT 233 271 WD 1.
 FT REPEAT 276 316 WD 2.
 FT REPEAT 339 379 WD 3.
 FT REPEAT 411 451 WD 4.
 SQ SEQUENCE 460 AA; 51022 MW; AA6F60448B7BCBA9 CRC64;

Query Match 64.3%; Score 36; DB 1; Length 460;
 Best Local Similarity 62.5%; Pred. No. 11;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VVPXGMHY 10
 :||| |:||
 Db 85 IVPLGLHY 92

RESULT 6

YK47_YEAST STANDARD; PRT; 743 AA.
 AC P36148;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Hypothetical 83.6 kDa protein in CCP1-MET1 intergenic region.
 GN YKR067W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RA van Vliet-Reedijk J.C., Planta R.J.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 CC -|- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -|- SIMILARITY: STRONG, TO YEAST YBL011W.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; Z28292; CAA82146.1; -;
 DR PIR; S38143; S38143.
 DR SGD; S0001775; YKR067W.
 DR InterPro; IPR002123; Acyltransferase.
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 31 55 POTENTIAL.
 FT TRANSMEM 69 85 POTENTIAL.
 FT TRANSMEM 502 524 POTENTIAL.
 FT TRANSMEM 539 555 POTENTIAL.
 SQ SEQUENCE 743 AA; 83644 MW; 84B9946E56B82F15 CRC64;

Query Match 64.3%; Score 36; DB 1; Length 743;
 Best Local Similarity 75.0%; Pred. No. 18;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VVPXGMHY 10
 :||| |:||
 Db 294 VVPCGLHY 301

RESULT 7

ID11_MESAU STANDARD; PRT; 227 AA.
 AC O35586;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Isopentenyl-diphosphate delta-isomerase 1 (EC 5.3.3.2) (IPP isomerase
 DE 1) (Isopentenyl pyrophosphate isomerase 1) (IPPI1).
 GN ID11.

OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Mesocricetus.
 OX NCBI_TaxID=10036;

RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=97373600; PubMed=9228075;
 RA Paton V.G., Shackelford J.E., Krisans S.K.;
 RT "Cloning and subcellular localization of hamster and rat isopentenyl
 RT diphosphate dimethylallyl diphosphate isomerase. A PTS1 motif targets
 RT the enzyme to peroxisomes".
 RL J. Biol. Chem. 272:18945-18950(1997).

CC -|- FUNCTION: CATALYZES THE 1,3-ALLYLIC REARRANGEMENT OF THE
 CC HOMOALLYLIC SUBSTRATE ISOPENTENYL (IPP) TO ITS HIGHLY
 CC ELECTROPHILIC ALLYLIC ISOMER, DIMETHYLLALLYL DIPHOSPHATE (DMAPP).
 CC -|- CATALYTIC ACTIVITY: Isopentenyl diphosphate = dimethylallyl
 CC diphosphate.

CC -|- COFACTOR: REQUIRES MAGNESIUM FOR ACTIVITY.

CC -|- PATHWAY: ISOPRENOID BIOSYNTHETIC PATHWAY WHOSE END PRODUCTS
 CC INCLUDE DOLICHOLS, VITAMINS A, D, E, AND K, STEROID HORMONES,
 CC CAROTENOIDS BILE ACIDS AND CHOLESTEROL.

CC -|- SUBCELLULAR LOCATION: Peroxisomal.

CC -|- SIMILARITY: BELONGS TO THE IPP ISOMERASE TYPE 1 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AF003836; AAC53283.1; -;
 DR InterPro; IPR002667; IPP isomerase.

DR InterPro; IPR000086; NUDIX_hydrolase.
 DR Pfam; PF00293; NUDIX; 1.

DR ProDom; PD004109; IPP_isomerase; 1.

KW isomerase; isoprene biosynthesis; Cholesterol biosynthesis;
 KW Sterol biosynthesis; Peroxisome; Magnesium.

FT ACT SITE 86 BY SIMILARITY.

FT ACT SITE 148 BY SIMILARITY.

FT SITE 225 227 MICROBODY TARGETING SIGNAL.

SQ SEQUENCE 227 AA; 26317 MW; F500A6586395E803 CRC64;

Query Match 62.5%; Score 35; DB 1; Length 227;
 Best Local Similarity 70.0%; Pred. No. 8.5;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMHY 10
 :||| |:||
 Db 121 EEVDPMEMHY 130

RESULT 8

GAAB_METTH STANDARD; PRT; 308 AA.
 AC O26806;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE GMP synthase [glutamine-hydrolyzing] subunit B (EC 6.3.5.2) (GMP
 DE synthetase).
 GN GAAB OR MTH710.

OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 OC Methanobacteriaceae; Methanothermobacter.
 OX NCBI_TaxID=187420;
 RN [1]
 RP SEQUENCE FROM N.A.

```
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwan N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155 (1997).
CC -!- CATALYTIC ACTIVITY: ATP + xanthosine 5'-phosphate + L-glutamine +
CC H(2)O = AMP + diphosphate + GMP + L-glutamate.
CC -!- PATHWAY: GMP biosynthesis.
CC -!- SUBUNIT: HETERODIMER COMPOSED OF A GLUTAMINE AMIDOTRANSFERASE
CC SUBUNIT (A) AND A GMP SYNTHASE SUBUNIT (B) (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE GMP SYNTHASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AE000850; RA085215.1; ALT_INIT.
DR HSSP; P04079; LGPM.
DR InterPro; IPR001674; GMP synt_C.
DR Pfam; PF00958; GMP synt_C; 1.
DR TIGRFAMs; TIGR00884; guaA Cterm; 1.
KW Ligase; GMP biosynthesis; Purine biosynthesis; ATP-binding;
KW Complete proteome.
FT DOMAIN 33 184 GMP-BINDING (BY SIMILARITY).
FT NP_BIND 29 35 ATP (BY SIMILARITY).
SQ SEQUENCE 308 AA; 34403 MW; F2DC6FD6202CAEC1 CRC64;
Query Match 62.5%; Score 35; DB 1; Length 308;
Best Local Similarity 63.6%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 EVVVPXGMHYS 11
DB 216 EWVESGLHES 226
RESULT 9
SLS1_YARLI
ID SLS1_YARLI STANDARD; PRT; 426 AA.
AC Q99158;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE SLS1 protein precursor.
GN SLS1.
OS Yarrowia lipolytica (Candida lipolytica).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=4952;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 20460 / W29;
RX MEDLINE=96216076; PubMed=8662639;
RA Boisrame A., Beckerich J.-M., Gaillardin C.;
RT "Sls1p, an endoplasmic reticulum component, is involved in the
RT protein translocation process in the yeast Yarrowia lipolytica.";
RL J. Biol. Chem. 271:11668-11675 (1996).
CC -!- FUNCTION: INVOLVED IN THE PROTEIN TRANSLOCATION PROCESS. MAY
CC INTERACT INDIRECTLY WITH TRANSLOCATING POLYPEPTIDES TO FACILITATE
CC THEIR TRANSFER AND/OR HELP THEIR FOLDING IN THE ER. IT IS NOT
CC REQUIRED FOR VIABILITY BUT IS ESSENTIAL FOR OPTIMAL GROWTH AT
CC ELEVATED TEMPERATURES.
CC
```

```
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; Z50154; CAA90516.1; --
DR InterPro; IPR000886; ER_target.
DR PROSITE; PS00014; ER_TARGET; 1.
KW Endoplasmic reticulum; Signal.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 426 SLS1 PROTEIN.
FT SITE 423 426 PREVENT SECRETION FROM ER (POTENTIAL).
SQ SEQUENCE 426 AA; 04CD7EF17540B9E2 CRC64;
Query Match 60.7%; Score 34; DB 1; Length 426;
Best Local Similarity 44.4%; Pred. No. 26;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 1 EVVVPXGMH 9
DB 52 DQVIPAGLH 60
RESULT 10
PHSL DESBA
ID PHSL_DESBA STANDARD; PRT; 513 AA.
AC P13085;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Periplasmic [NiFeSe] hydrogenase large subunit (SC 1.18.99.1) (NiFeSe
DE hydrogenlyase large chain).
OS Desulfovibrio baculatus (Desulfomicrobium baculatus).
OC Bacteria; Proteobacteria; delta subdivision; Desulfomicrobium.
OX NCBI_TaxID=899;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88058744; PubMed=3316183;
RA Menon N.K., Peck H.D. Jr., le Gall J., Przybyla A.E.;
RT "Cloning and sequencing of the genes encoding the large and small
RT subunits of the periplasmic (NiFeSe) hydrogenase of Desulfovibrio
RT baculatus.";
RL J. Bacteriol. 169:5401-5407 (1987).
RN [2]
RP REVISIONS.
RA Menon N.K., Peck H.D. Jr., le Gall J., Przybyla A.E.;
RL J. Bacteriol. 170:4429-4429 (1988).
RN [3]
RX X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS).
RX MEDLINE=99306038; PubMed=10378275;
RA Garcin E., Vernece X., Hatchikian E.C., Volbeda A., Frey M.,
RA Fontecilla-Camps J.C.;
RT "The crystal structure of a reduced [NiFeSe] hydrogenase provides an
RT image of the activated catalytic center.";
RL Structure 7:557-566 (1999).
CC -!- CATALYTIC ACTIVITY: 2 reduced ferredoxin + 2 H(+) = 2 oxidized
CC ferredoxin + H(2).
CC -!- COFACTOR: ONE NICKEL ION, TWO 4FE-4S CLUSTERS AND ONE
CC SELENOCYSTEINE.
CC -!- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL SUBUNIT.
CC -!- SUBCELLULAR LOCATION: Periplasmic.
CC -!- MISCELLANEOUS: PERHAPS THE LEADER OF THE SMALL SUBUNIT SERVES AS A
CC TRANSPORT VEHICLE FOR BOTH SUBUNITS.
CC -!- SIMILARITY: BELONGS TO THE [NIFE] / [NIFESE] HYDROGENASE LARGE
CC SUBUNIT FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M18271; AAA23375.2; -
 PIR; A33101; HQDVLB.
 PDB; 1CC1; 01-JUN-99.
 InterPro: IPR001501; Ni_hdl.
 Pfam: PF00374; NiFeSe_Hases; 1.
 PROSITE; PS00507; NI_HGENASE_L1; 1.
 PROSITE; PS00508; NI_HGENASE_L2; 1.
 Oxidoreductase; Periplasmic; Nickel; Selenium; Selenocysteine;
 3D-structure.
 INIT_MET 0 0
 METAL 51 51 IRON 2.
 METAL 70 70 NICKEL.
 METAL 73 73 IRON 1 AND NICKEL.
 METAL 444 444 IRON 2.
 METAL 492 492 NICKEL.
 METAL 495 495 IRON 1 AND NICKEL.
 METAL 498 498 IRON 2.
 SE_CYS 492 492
 SQ SEQUENCE 513 AA; 56683 MW; AC9285A6F80576FC CRC64;

Query Match 58.9%; Score 33; DB 1; Length 513;
 Best Local Similarity 71.4%; Pred. No. 50;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 5 PXGMHVS 11
 | | | | |
 Db 297 PGGLHVS 303

RESULT 11

MUTL_BACSU STANDARD; PRT; 627 AA.
 AC P49850;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA mismatch repair protein mutL.
 GN MUTL.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=96349107; PubMed=8760914;
 RA Ginetti F., Perego M., Albertini A.M., Galizzi A.;
 RT "Bacillus subtilis mutS mutL operon: identification, nucleotide
 sequence and mutagenesis.";
 RL Microbiology 142:2021-2029(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 Brouillet S., Bruschini C.V., Caldwell B., Capuano V., Carter N.M.,
 Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
 Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 Etian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 Ghim S.Y., Glaser P., Goffey A., Gollightly E.J., Grandi G.,
 Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut J.,
 Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 Kobayashi Y., Koester P., Konigstein G., Krogh S., Kumano M.,
 Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,

Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
 Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 Rieger M., Rivoita C., Roche E., Roche M., Rose M., Sadale Y.,
 Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
 Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 Viari A., Wambut R., Wedler E., Wedler H., Weitzenecker T.,
 Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*";
 RL Nature 390:249-256(1997).
 CC -!- FUNCTION: THIS PROTEIN IS INVOLVED IN THE REPAIR OF MISMATCHES IN
 DNA. IT IS REQUIRED FOR DAM-DEPENDENT METHYL-DIRECTED DNA MISMATCH
 REPAIR. MAY ACT AS A "MOLECULAR MATCHMAKER", A PROTEIN THAT
 PROMOTES THE FORMATION OF A STABLE COMPLEX BETWEEN TWO OR MORE
 DNA-BINDING PROTEINS IN AN ATP-DEPENDENT MANNER WITHOUT ITSELF
 BEING PART OF A FINAL EFFECTOR COMPLEX (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTL/HEXB FAMILY.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; U27343; AAB19236.1; -
 EMBL; Z99112; CAB13578.1; -
 HSP; P23367; IBKN.
 DR Subtilisin; EGI1402; mutL.
 DR InterPro; IPR003594; ATPbind ATPase.
 DR InterPro; IPR002099; DNA mismatch repair.
 DR InterPro; IPR004359; HIS_KIN_sig.
 DR Pfam; PF01119; DNA mismatch repair; 1.
 DR Pfam; PF02518; HATPase_C; 1.
 DR TIGRFAMs; TIGR00585; mutL; 1.
 DR PROSITE; PS00058; DNA MISMATCH_REPAIR_1; 1.
 DR DNA repair; Complete proteome.

Query Match 58.9%; Score 33; DB 1; Length 627;
 Best Local Similarity 54.5%; Pred. No. 61;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 BEVVPXGMHVS 11
 | | | | |
 Db 488 EMIVPLTFHVS 498

RESULT 12

KPBA CAEEL STANDARD; PRT; 1188 AA.
 AC P3435;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable phosphorylase B kinase alpha regulatory chain (Phosphorylase
 kinase alpha subunit).
 GN C14B9.8.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;

```

RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonnenhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP REVISIONS.
RA Waterston R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PHOSPHORYLASE B KINASE CATALYZES THE PHOSPHORYLATION OF
CC SERINE IN CERTAIN SUBSTRATES, INCLUDING TROPONIN I. THE ALPHA
CC CHAIN MAY BIND CALMODULIN (BY SIMILARITY).
CC -!- PATHWAY: Glycogen metabolism.
CC -!- SIMILARITY: BELONGS TO THE PHOSPHORYLASE B KINASE REGULATORY
CC CHAINS FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; L15188; AA27954.2; -
DR PIR; S44754; S44754.
DR WormPep; C14B9.8; CE26870.
KW Hypothetical protein; Glycogen metabolism; Calmodulin-binding.
SQ SEQUENCE 1188 AA; 135558 MW; DE9BB875F3603863 CRC64;

Query Match 58.9%; Score 33; DB 1; Length 1188;
Best Local Similarity 54.5%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 BEVFXGMHYS 11
|||
|.|
Db 950 EYAEEDGHYS 960

RESULT 13
VLTF_BPT5
ID VLTF_BPT5 STANDARD; PRT; 1396 AA.
AC P13390; O48502; PRT; 1396 AA.
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE L-shaped tail fiber protein (LTF protein).
GN LTF.
OS Bacteriophage T5.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC T5-like viruses.
OC NCBI_TaxID=10726;
OX [1]
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95309401; PubMed=7789514;
RA Kallman A.V., Kulshin V.E., Shiyapnikov M.G., Ksenzenko V.N.,
RA Kryukov V.M.;
RT "The nucleotide sequence of the bacteriophage T5 ltf gene.";
RL FEBS Lett. 366:46-48(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Kallman A.V.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.

```

```

RN [3]
RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
RX MEDLINE=88289370; PubMed=3267228;
RA Kallman A.V., Kryukov V.M., Bayev A.A.;
RT "The nucleotide sequence of bacteriophage T5 DNA at the region
RT between early and late genes.";
RL Nucleic Acids Res. 16:6230-6230(1988).
CC -!- FUNCTION: NONESSENTIAL PROTEIN THAT MEDIATES BINDING TO THE
CC POLYMANNOSE O ANTIGEN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X69460; CAA49220.1; -
DR EMBL; AJ001191; CAA04591.1; -
DR PIR; S01982; S01982.
KW Late protein.
FT CONFLICT 986 986 V -> A (IN REF. 2).
SQ SEQUENCE 1396 AA; 147989 MW; 18CD2192F65FFFC1 CRC64;

Query Match 58.9%; Score 33; DB 1; Length 1396;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVFXGMHYS 11
:|:|
Db 1360 KTIPAGDHYS 1369

RESULT 14
TKNK_BOVIN
ID TKNK_BOVIN STANDARD; PRT; 126 AA.
AC P08858;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurokinin B precursor (NKB) (Neuromedin K).
GN TAC3 OR NKKB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86313713; PubMed=3462746;
RA Kotani H., Hoshimaru M., Nawa H., Nakanishi S.;
RT "Structure and gene organization of bovine neuromedin K precursor.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:7074-7078(1986).
CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC MUSCLES.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M14351; AAA30723.1; -
DR EMBL; M14347; AAA30723.1; JOINED.
DR EMBL; M14348; AAA30723.1; JOINED.
DR EMBL; M14349; AAA30723.1; JOINED.

```

DR EMBL; M14350; AAA30723.1; JOINED.
DR PIR; A25905; A25905.
DR InterPro; IPR003635; Neurokinin.
DR InterPro; IPR002040; Tachykinin.
DR ProDom; PD020370; Neurokinin; 1.
DR PROSITE; PS00267; TACHYKININ; 1.
KW Tachykinin; Neuropeptide; Cleavage on pair of basic residues;
KW Amidation; Signal.

FT SIGNAL 1 20 POTENTIAL.
FT PROPEP 21 83
FT PEPTIDE 86 95 NEUROKININ B.
FT PROPEP 99 126
FT MOD_RES 95 95
SQ SEQUENCE 126 AA; 13871 MW; 446EF433498EC059 CRC64;

Query Match 57.1%; Score 32; DB 1; Length 126;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMH 9
Db 28 EQVVPGGH 36

RESULT 15

RR2_CHLVU ID_RR2_CHLVU STANDARD; PRT; 267 AA.
AC P56351;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Chloroplast 30S ribosomal protein S2.
GN RPS2.
OS Chlorella vulgaris.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
OC Chlorellaceae; Chlorella.
OX NCBI_TaxID=3077;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IAM C-27 / Tamiya;
RX MEDLINE=97303241; PubMed=9159184;
RA Wakasugi T., Nagai T., Kapoor M., Sugita M., Ito M., Ito S.,
RA Tsudzuki J., Nakashima K., Tsudzuki T., Suzuki Y., Hamada A., Ohta T.,
RA Inamura A., Yoshinaga K., Sugiura M.;
RT "Complete nucleotide sequence of the chloroplast genome from the
green alga Chlorella vulgaris: the existence of genes possibly
involved in chloroplast division".
RL Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).
CC -!- SIMILARITY: BELONGS TO THE S2P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AB001684; BAA57862.1; -
DR InterPro; IPR001865; Ribosomal_S2.
DR Pfam; PF00318; Ribosomal_S2; 1.
DR PRINTS; PR00395; RIBOSOMALS2.
DR TIGRFAMS; TIGR01011; rpsB_bact; 1.
DR PROSITE; PS00962; RIBOSOMAL_S2_1; FALSE_NEG.
DR PROSITE; PS00963; RIBOSOMAL_S2_2; 1.
KW Ribosomal protein; Chloroplast.
SQ SEQUENCE 267 AA; 30699 MW; 7903075340BD900F CRC64;

Query Match 57.1%; Score 32; DB 1; Length 267;
Best Local Similarity 40.0%; Pred. No. 40;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10
Db 8 EDMIQSGMHF 17

Search completed: June 4, 2003, 13:11:45
Job time : 7.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:08:49 ; Search time 11 Seconds
(without alignments)
96.134 Million cell updates/sec

Title: AUDET-909-2

Perfect score: 56

Sequence: 1 eevvpxgmhys 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	66.1	308	2 A72807	ftsH proteinase ac
2	37	66.1	1057	2 F89892	carbamoyl-phosphat
3	36	64.3	102	2 A42452	V1 protein - tobac
4	36	64.3	252	2 AE2001	hypothetical prote
5	36	64.3	460	2 S69046	hypothetical prote
6	36	64.3	743	2 S38143	hypothetical prote
7	35	62.5	156	2 D82618	conserved hypotet
8	35	62.5	233	2 T02590	DNA binding protei
9	35	62.5	311	2 H69194	GMP synthetase, su
10	35	62.5	425	2 T24111	hypothetical prote
11	35	62.5	510	2 G86430	T518.1 protein - A
12	34	60.7	264	2 G69117	thiophene synthase
13	34	60.7	279	2 C75538	hypothetical prote
14	34	60.7	350	2 B75478	3-dehydroquinatase
15	34	60.7	355	2 T35025	probable DNA ligase
16	34	60.7	360	2 E69086	cell division prot
17	34	60.7	425	2 C83903	hypothetical prote
18	34	60.7	426	2 S58132	hypothetical prote
19	34	60.7	495	2 T28717	hypothetical prote
20	34	60.7	1028	2 AF3286	ATP-dependent DNA
21	33	58.9	156	2 S54619	hypothetical prote
22	33	58.9	367	2 E83607	polyamine transpor
23	33	58.9	441	2 G82253	conserved hypotet
24	33	58.9	466	2 G71542	probable amino aci
25	33	58.9	466	2 H81697	amino acid antipor
26	33	58.9	487	2 S65811	finger protein (cl
27	33	58.9	514	1 HQDVLB	cytochrome-c3 hydr
28	33	58.9	534	2 A63284	coenzyme F420-quin
29	33	58.9	545	2 T08564	hypothetical prote

30 33 58.9 627 2 A69663 DNA mismatch repair
31 33 58.9 716 1 JCS061 macrophage-stimula
32 33 58.9 1257 2 S44754 C14B9.8 protein -
33 33 58.9 1396 2 S36851 L-shaped tail fibe
34 33 58.9 3472 2 T31308 hypothetical tail 367K
35 32 57.1 126 2 A25905 tachykinin B precu
36 32 57.1 197 2 D71640 heme exporter prot
37 32 57.1 225 2 S57810 hypothetical prote
38 32 57.1 233 2 E97120 ribosomal protein
39 32 57.1 267 2 T07215 ribosomal protein
40 32 57.1 270 2 C95881 probable ThuaA prot
41 32 57.1 283 2 T25737 hypothetical prote
42 32 57.1 288 2 JC4011 cyclin D2 - rat
43 32 57.1 288 2 I58372 cyclin D2 - mouse
44 32 57.1 289 2 A41984 cyclin D2 - human
45 32 57.1 289 2 A42822 cyclin D2 - human

ALIGNMENTS

RESULT 1

A72207

ftsH proteinase activity modulator HflK - Thermotoga maritima (strain MSB8)

C;Species: Thermotoga maritima

C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000

C;Accession: A72207

R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.

Nature 398, 323-329, 1999

A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq

A;Reference number: A72200; MUID:99287316; PMID:10360571

A;Accession: A72207

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-308 <ARN>

A;Cross-references: GB:AE001819; GB:AE000512; NID:g4982396; PIDN:AAD36885.1; PID:g4982404

A;Experimental source: strain MSB8

C;Genetics:

A;Gene: TW1822

C;Superfamily: erythrocyte band 7 integral membrane protein

Query Match 66.1%; Score 37; DB 2; Length 308;

Best Local Similarity 75.0%; Pred. No. 10;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMHY 10

DB 41 VVPSGIHY 48

RESULT 2

F89892

carbamoyl-phosphate synthase large chain [imported] - Staphylococcus aureus (strain N315)

C;Species: Staphylococcus aureus

C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C;Accession: F89892

R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguchi, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A;Reference number: A89758; MUID:21311952; PMID:11418146

A;Accession: F89892

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1057 <KUR>

A;Cross-references: GB:BA000018; PID:gl3701002; PIDN:BA842298.1; GSPDB:GN00149

A;Experimental source: strain N315

C;Genetics:

A;Gene: pyrAB

C;Superfamily: carbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain; biotin

Query Match 66.1%; Score 37; DB 2; Length 1057;
Best Local Similarity 60.0%; Pred. No. 39;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMHYS 11
|:| |:
Db 190 EIVSNGLHYS 199

RESULT 3
A42452
V1 protein - tobacco yellow dwarf virus (strain Australia)
C:Species: tobacco yellow dwarf virus
C:Date: 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change 08-Oct-1999
C:Accession: A42452
R:Morris, B.A.M.; Richardson, K.A.; Haley, A.; Zhan, X.; Thomas, J.E.
Virology 187, 633-642, 1992
A:Title: The nucleotide sequence of the infectious cloned DNA component of tobacco yellow
A:Reference number: A42452; MUID:92188538; PMID:1546458
A:Accession: A42452
A:Molecule type: DNA
A:Residues: 1-102 <MOR>
A:Cross-references: GB:M81103; NID:G335283; PIDN:AAA47947.1; PID:G335284

Query Match 64.3%; Score 36; DB 2; Length 102;
Best Local Similarity 60.0%; Pred. No. 5;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMHYS 11
|:| |:
Db 7 QVVPISGINS 16

RESULT 4
AE2001
hypothetical protein alr1563 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AE2001
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AE2001
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-252 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA077929.1; PID:G17135383; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr1563

Query Match 64.3%; Score 36; DB 2; Length 252;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10
|:| |:
Db 235 EMIVPAGLHF 244

RESULT 5
S69046
hypothetical protein YPL139c - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C:Date: 22-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 17-Mar-2000
C:Accession: S69046
R:Hall, J.; DePaulo, T.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Storms, R.K.;
submitted to the EMBL Data Library, December 1995
A:Description: The sequence of Saccharomyces cerevisiae chromosome XVI left arm.

A:Reference number: S69040
A:Accession: S69046
A:Molecule type: DNA
A:Residues: 1-460 <HAL>
A:Cross-references: EMBL:U43703; NID:G1244769; PIDN:AA868221.1; PID:G1244776; MIPS:YPL139c
C:Genetics:
A:Gene: SGD:UME1
A:Cross-references: SGD:S0006060; MIPS:YPL139c
A:Map position: 161
C:Superfamily: Saccharomyces cerevisiae transcription modulator WTMI

Query Match 64.3%; Score 36; DB 2; Length 460;
Best Local Similarity 62.5%; Pred. No. 25;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMHY 10
|:| |:
Db 85 IVPLGLHY 92

RESULT 6
S38143
hypothetical protein YBL011w homolog YKR067w - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 19-Apr-2002
A:Accession: S38143
R:van Vliet-Reedijk, J.C.; Planta, R.J.
submitted to the Protein Sequence Database, March 1994
A:Reference number: S38130
A:Accession: S38143
A:Molecule type: DNA
A:Residues: 1-743 <VAN>
A:Cross-references: EMBL:Z28292; NID:G486536; PIDN:CRA82146.1; PID:G486537; MIPS:YKR067w
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:GPT2
A:Cross-references: SGD:S0001775
A:Map position: 11R
C:Keywords: transmembrane protein

Query Match 64.3%; Score 36; DB 2; Length 743;
Best Local Similarity 75.0%; Pred. No. 43;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMHY 10
|:| |:
Db 294 VVPGLHY 301

RESULT 7
D82618
conserved hypothetical protein XFI950 [imported] - Xylella fastidiosa (strain 985c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: D82618
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequenc
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: D82618
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-156 <SIM>
A:Cross-references: GB:AE004014; GB:AE003849; NID:G9107044; PIDN:AAF84752.1; GSPDB:GN0017
A:Experimental source: strain 985c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Al
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, H.
as-Neto, E.; Docena, C.; El-Dorriy, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Praga, J.S.; Franca, S.C.; Franco, M.C.; Frohme
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; Laigre
chado, M.A.; Madeira, A.M.B.N.; Marino, C.L.; Marques, M.V.; Martins, E.

A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1950

Query Match 62.5%; Score 35; DB 2; Length 156;
Best Local Similarity 55.6%; Pred. No. 13;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 BEVVPXGMHY 9

Db 119 BEILPQGVH 127

RESULT 8

T02590

DNA binding protein EREBP-2 - common tobacco

C;Species: Nicotiana tabacum (common tobacco)

C;Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 21-Jul-2000

C;Accession: T02590

R;Ohme-Takagi, M.; Shinshi, H.

Plant Cell 7, 173-182, 1995

A;Title: Ethylene-inducible DNA binding proteins that interact with an ethylene responsi

A;Reference number: Z14671; MUID:95276459; PMID:7756828

A;Accession: T02590

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-233 <OHM>

A;Cross-references: EMBL:D38126; NID:G790362; PIDN:BA07324.1; PID:g1208498

A;Experimental source: strain BV4; tissue-type leaf

QY 1 BEVVPXGMHY 10

Db 90 QAVVPKGRHY 99

Query Match 62.5%; Score 35; DB 2; Length 233;
Best Local Similarity 60.0%; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

RESULT 9

H69194

GMP synthetase, subunit B - Methanobacterium thermoautotrophicum (strain Delta H)

C;Species: Methanobacterium thermoautotrophicum

C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999

C;Accession: H69194

R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;

Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.

Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func

A;Reference number: A59000; MUID:98037514; PMID:9371463

A;Accession: H69194

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-311 <MTH>

A;Cross-references: GB:AE000850; GB:AE000666; NID:G2621794; PIDN:AAB85215.1; PID:g262179

A;Experimental source: strain Delta H

C;Genetics:

A;Gene: MTH10

A;Start codon: GTG

Query Match 62.5%; Score 35; DB 2; Length 311;
Best Local Similarity 63.6%; Pred. No. 27;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 BEVVPXGMHY 11

Db 119 BEILPQGVH 127

Db 219 EEVVEGLHES 229

RESULT 10

T24111

hypothetical protein R10D12.10 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C;Accession: T24111

R;Percy, C.

submitted to the EMBL Data Library, October 1996

A;Reference number: Z19842

A;Accession: T24111

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-425 <WIL>

A;Cross-references: EMBL:Z81109; PIDN:CAB03241.1; GSPDB:GN00023; CESP:R10D12.10

A;Experimental source: clone R10D12

C;Genetics:

A;Gene: CESP:R10D12.10

A;Map position: 5

A;Introns: 23/3; 56/3; 113/3; 257/2

Query Match 62.5%; Score 35; DB 2; Length 425;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 BEVVPXGMHY 10

Db 335 EQIVPGGLQY 344

RESULT 11

G86430

TS18.1 protein - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cross)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 17-May-2002

C;Accession: G86430

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: G86430

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-510 <STO>

A;Cross-references: GB:AE005172; NID:G4587512; PIDN:AAD25743.1; GSPDB:GN00141

C;Genetics:

A;Map position: 1

C;Superfamily: hexose phosphate transport protein uhpt

Query Match 62.5%; Score 35; DB 2; Length 510;
Best Local Similarity 60.0%; Pred. No. 45;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10

Db 12 EEVVPGLHF 21

RESULT 12

G69117

diphthine synthase - Methanobacterium thermoautotrophicum (strain Delta H)

C;Species: Methanobacterium thermoautotrophicum

C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 08-Oct-1999

C;Accession: G69117

R.; Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني, N.; Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7135, 1997.
 A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
 A;Reference number: A69000; MUID:98037514; PMID:93711463
 A;Accession: G69117
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-264 <WTH>
 A;Cross-references: GB:AE000940; GB:AE000666; NID:G2623011; PIDN:AAB86340.1; PID:G2623011
 A;Experimental source: strain Delta H
 C;Genetics:
 A;Gene: MTH1874
 C;Superfamily: Methanococcus jannaschii diphthine synthase

Query Match 60.7%; Score 34; DB 2; Length 264;
 Best Local Similarity 62.5%; Pred. No. 35;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMHY 10
 |||||
 Db 235 VVPAGLHF 242
 |||||

RESULT 13
 C75538
 hypothetical protein - Deinococcus radiodurans (strain R1)
 C;Species: Deinococcus radiodurans
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 28-Jul-2000
 C;Accession: C75538
 R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999

A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A;Reference number: A75250; MUID:20036896; PMID:10567266
 A;Accession: C75538
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-279 <WHI>
 A;Cross-references: GB:AE001889; GB:AE000513; NID:G6457944; PIDN:AAF09867.1; PID:G645795
 A;Experimental source: strain R1
 C;Genetics:
 A;Gene: DR0271
 A;Map position: 1
 C;Superfamily: Deinococcus radiodurans hypothetical protein DR0271

Query Match 60.7%; Score 34; DB 2; Length 279;
 Best Local Similarity 75.0%; Pred. No. 38;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 VPXGMHY 11
 |||||
 Db 100 VPLGRHY 107
 |||||

RESULT 14
 B75478
 3-dehydroquininate synthase - Deinococcus radiodurans (strain R1)
 C;Species: Deinococcus radiodurans
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
 C;Accession: B75478
 R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999

A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A;Reference number: A75250; MUID:20036896; PMID:10567266
 A;Accession: B75478
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-350 <WHI>

Query Match 60.7%; Score 34; DB 2; Length 350;
 Best Local Similarity 60.0%; Pred. No. 48;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPGMHYS 11
 |||||
 Db 252 EAVAYGMHYA 261
 |||||

RESULT 15
 T35025
 probable DNA ligase - Streptomyces coelicolor
 C;Species: Streptomyces coelicolor
 C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
 C;Accession: T35025
 R;Seeger, S.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, June 1999
 A;Reference number: Z21565
 A;Accession: T35025
 A;Status: preliminary; translated from GB/EMBL/DBDJ
 A;Molecule type: DNA
 A;Residues: 1-355 <SEE>
 A;Cross-references: EMBL:AL079355; PIDN:CAB45581.1; GSPDB:GN00070; SCOEDB:SC4C6.17c
 A;Experimental source: strain A3(2)
 C;Genetics:
 A;Gene: SCOEDB:SC4C6.17c

Query Match 60.7%; Score 34; DB 2; Length 355;
 Best Local Similarity 71.4%; Pred. No. 49;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VPXGMHY 10
 :|||
 Db 20 IPPGMHY 26
 :|||

A;Cross-references: GB:AE001932; GB:AE000513; NID:G6458481; PIDN:AAF10353.1; PID:G6458481
 A;Experimental source: strain R1
 C;Genetics:
 A;Gene: DR0777
 A;Map position: 1
 C;Superfamily: 3-dehydroquininate synthase; 3-dehydroquininate synthase homology

Query Match 60.7%; Score 34; DB 2; Length 350;
 Best Local Similarity 60.0%; Pred. No. 48;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPGMHYS 11
 |||||
 Db 252 EAVAYGMHYA 261
 |||||

RESULT 15
 T35025
 probable DNA ligase - Streptomyces coelicolor
 C;Species: Streptomyces coelicolor
 C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
 C;Accession: T35025
 R;Seeger, S.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, June 1999
 A;Reference number: Z21565
 A;Accession: T35025
 A;Status: preliminary; translated from GB/EMBL/DBDJ
 A;Molecule type: DNA
 A;Residues: 1-355 <SEE>
 A;Cross-references: EMBL:AL079355; PIDN:CAB45581.1; GSPDB:GN00070; SCOEDB:SC4C6.17c
 A;Experimental source: strain A3(2)
 C;Genetics:
 A;Gene: SCOEDB:SC4C6.17c

Query Match 60.7%; Score 34; DB 2; Length 355;
 Best Local Similarity 71.4%; Pred. No. 49;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VPXGMHY 10
 :|||
 Db 20 IPPGMHY 26
 :|||

Search completed: June 4, 2003, 13:14:14
 Job time : 12 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:09:19 ; Search time 9.75 Seconds
(without alignments)
33.195 Million cell updates/sec

Title: AUDET-909-2

Perfect score: 56

Sequence: 1 eevvpqgmhys 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/ECTUS-COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	69.6	1037	4	US-09-134-001C-4794
2	34	60.7	600	2	US-08-821-119-19
3	34	60.7	600	2	US-08-821-118-2
4	33	58.9	747	4	US-09-724-864-36
5	32	57.1	70	4	US-09-134-001C-3950
6	32	57.1	102	2	US-08-580-988A-23
7	32	57.1	126	2	US-08-879-995A-3
8	32	57.1	126	3	US-09-215-096-3
9	32	57.1	152	2	US-08-460-694-4
10	32	57.1	152	3	US-08-460-744-4
11	32	57.1	152	3	US-07-667-711B-4
12	32	57.1	173	1	US-08-193-977-7
13	32	57.1	189	2	US-08-464-517-21
14	32	57.1	189	2	US-08-246-361A-21
15	32	57.1	189	3	US-08-463-772-21
16	32	57.1	189	5	PCT-US93-05000-21
17	32	57.1	236	2	US-08-464-517-22
18	32	57.1	236	2	US-08-246-361A-22
19	32	57.1	236	3	US-08-463-772-22
20	32	57.1	236	5	PCT-US93-05000-22
21	32	57.1	280	2	US-08-464-517-6
22	32	57.1	280	3	US-08-463-772-6
23	32	57.1	289	2	US-08-246-361A-4
24	32	57.1	289	5	PCT-US93-05000-4
25	32	57.1	291	5	PCT-US93-05000-6
26	32	57.1	292	2	US-08-464-517-23
27	32	57.1	292	2	US-08-246-361A-6

```

28 32 57.1 292 2 US-08-246-361A-23 Sequence 23, Appl
29 32 57.1 292 3 US-08-463-772-23 Sequence 23, Appl
30 32 57.1 292 5 PCT-US93-05000-23 Sequence 23, Appl
31 32 57.1 295 1 US-07-947-120-8 Sequence 8, Appl
32 32 57.1 295 1 US-08-472-893A-8 Sequence 8, Appl
33 32 57.1 295 2 US-08-460-694-2 Sequence 2, Appl
34 32 57.1 295 2 US-08-464-517-19 Sequence 19, Appl
35 32 57.1 295 2 US-08-464-517-20 Sequence 20, Appl
36 32 57.1 295 2 US-08-246-361A-19 Sequence 19, Appl
37 32 57.1 295 2 US-08-246-361A-20 Sequence 20, Appl
38 32 57.1 295 3 US-08-463-772-19 Sequence 19, Appl
39 32 57.1 295 3 US-08-463-772-20 Sequence 20, Appl
40 32 57.1 295 3 US-08-460-744-2 Sequence 2, Appl
41 32 57.1 295 3 US-07-667-711B-2 Sequence 2, Appl
42 32 57.1 295 3 US-08-947-492-8 Sequence 8, Appl
43 32 57.1 295 5 PCT-US93-05000-2 Sequence 19, Appl
44 32 57.1 295 5 PCT-US93-05000-19 Sequence 20, Appl
45 32 57.1 295 5 PCT-US93-05000-20 Sequence 20, Appl

```

ALIGNMENTS

RESULT 1

US-09-134-001C-4794
; Sequence 4794, Application US/09134001C
; Patent No. 6380370

GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GPC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4794

; LENGTH: 1037
; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4794

Query Match 69.6%; Score 39; DB 4; Length 1037;
Best Local Similarity 63.8%; Pred. No. 16;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EEVVPXGMHYS 11

Db 199 KEVVSNGHYS 209

RESULT 2

US-08-821-119-19

; Sequence 19, Application US/08821119
; Patent No. 5821104

GENERAL INFORMATION:

; APPLICANT: Holm, Kaj Andre

; APPLICANT: Rasmussen, Grethe

; APPLICANT: Halkier, Torben

; APPLICANT: Lehmbeck, Jan

; TITLE OF INVENTION: Tripeptidyl Aminopeptidase

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 5821104o No. 5821104disk of No. 5821104th America, Inc.

; STREET: 405 Lexington Avenue

; CITY: New York

; STATE: NY

; COUNTRY: USA

; ZIP: 10174

; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/821,119
; FILING DATE: 19-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4107.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:

```

INFORMATION FOR SEQ ID NO: 19:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 600 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-821-119-19

```

```

Query Match      60.7%; Score 34; DB 2; Length 600;
Best Local Similarity 75.0%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      4 VPXGMHYS 11
      |||||
DB      31 VPKGWHYS 38

```

RESULT 3

```

US-08-821-118-2
; Sequence 2, Application US/08821118
; Patent No. 5989889

```

GENERAL INFORMATION:

```

; APPLICANT: Rey, Michael
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING POLYPEPTIDES
; TITLE OF INVENTION: HAVING TRIPEPTIDE AMINOPEPTIDASE
; TITLE OF INVENTION: ACTIVITY
; NUMBER OF SEQUENCES: 7

```

CORRESPONDENCE ADDRESS:

```

; ADDRESSEE: No. 59898890 No. 5989889disk of No. 5989889th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA

```

```

; ZIP: 10174

```

COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/821,118

```

```

; FILING DATE: 19-MAR-1997

```

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

```

; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4107.400-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:

```

INFORMATION FOR SEQ ID NO: 2:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 600 amino acids

```

```

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-821-118-2

```

```

Query Match      60.7%; Score 34; DB 2; Length 600;
Best Local Similarity 75.0%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      4 VPXGMHYS 11
      |||||
DB      31 VPKGWHYS 38

```

RESULT 4

```

US-09-724-864-36
; Sequence 36, Application US/09724864
; Patent No. 6380362

```

GENERAL INFORMATION:

```

; APPLICANT: Watson, James D.
; APPLICANT: Murison, James G.
; TITLE OF INVENTION: Polynucleotides, polypeptides expressed
; TITLE OF INVENTION: by the polynucleotides and methods for their use.
; FILE REFERENCE: 11000.1050U1

```

```

; CURRENT APPLICATION NUMBER: US/09/724,864

```

```

; CURRENT FILING DATE: 2000-11-28

```

```

; PRIOR APPLICATION NUMBER: U.S. No. 6380362 60/171,678

```

```

; PRIOR FILING DATE: 1999-12-23

```

```

; NUMBER OF SEQ ID NOS: 72

```

```

; SOFTWARE: FastSEQ for Windows Version 4.0

```

```

; SEQ ID NO 36

```

```

; LENGTH: 747

```

```

; TYPE: PRT

```

```

; ORGANISM: Rat

```

```

; US-09-724-864-36

```

```

Query Match      58.9%; Score 33; DB 4; Length 747;
Best Local Similarity 71.4%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      5 PXGMHYS 11
      |||||
DB      627 PGLHYS 633

```

RESULT 5

```

US-09-134-001C-3950

```

```

; Sequence 3950, Application US/09134001C
; Patent No. 6380370

```

GENERAL INFORMATION:

```

; APPLICANT: Lynn Doucette-Stamm et al

```

```

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

```

```

; FILE REFERENCE: GTC-007

```

```

; CURRENT APPLICATION NUMBER: US/09/134,001C

```

```

; CURRENT FILING DATE: 1998-08-13

```

```

; PRIOR APPLICATION NUMBER: US 60/064,964

```

```

; PRIOR FILING DATE: 1997-11-08

```

```

; PRIOR APPLICATION NUMBER: US 60/055,779

```

```

; PRIOR FILING DATE: 1997-08-14

```

```

; NUMBER OF SEQ ID NOS: 5674

```

```

; SEQ ID NO 3950

```

```

; LENGTH: 70

```

```

; TYPE: PRT

```

```

; ORGANISM: Staphylococcus epidermidis

```

```

; US-09-134-001C-3950

```

```

Query Match      57.1%; Score 32; DB 4; Length 70;
Best Local Similarity 62.5%; Pred. No. 19;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

QY 4 VPXGMHYS 11
:|||||
Db 36 MPKGFHYS 43

RESULT 6

US-08-580-988A-23
; Sequence 23, Application US/08580988A
; Patent No. 5956161
; GENERAL INFORMATION:
; APPLICANT: Aggarwal et al.
; TITLE OF INVENTION: Tumor Necrosis Factor
; TITLE OF INVENTION: Receptor-I-Associated Protein Kinase And Methods
; TITLE OF INVENTION: For Its Use
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. Benjamin A. Adler
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/580,988A
; FILING DATE: January 3, 1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benjamin Aaron Adler, Ph.D., J.D.
; REGISTRATION NUMBER: 35,423
; REFERENCE/DOCKET NUMBER: D5721CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-777-2321
; TELEFAX: 713-777-6908
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 102 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: protein
; HYPOTHETICAL: no
; ANTI-SENSE: no
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
US-08-580-988A-23

Query Match 57.1%; Score 32; DB 2; Length 102;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 BEVVPXGMHY 10
|||||:
Db 24 BEVFPAMNY 33

RESULT 7

US-08-879-995A-3
; Sequence 3, Application US/08879995A
; Patent No. 5985606
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: HUMAN PREPROTACHYKININ B

; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/879,995A
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PP-0326 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 163590
US-08-879-995A-3

Query Match 57.1%; Score 32; DB 2; Length 126;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 BEVVPXGMH 9
:|||||
Db 28 EQVWPGGH 36

RESULT 8

US-09-215-096-3
; Sequence 3, Application US/09215096
; Patent No. 6008194
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: HUMAN PREPROTACHYKININ B
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/215,096
; FILING DATE:

```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/879,995
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0326 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 163590
; US-09-215-096-3

```

```

Query Match 57.1%; Score 32; DB 3; Length 126;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 EEVVPXGMH 9
Db 28 EQVPGGHH 36

```

```

RESULT 9
US-08-460-694-4
; Sequence 4, Application US/08460694
; Patent No. 5858655
; GENERAL INFORMATION:
; APPLICANT: Arnold, Andrew
; TITLE OF INVENTION: Prad1 Cyclin and its cDNA
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,694
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McConathy, Evelyn H.
; REGISTRATION NUMBER: 35,279
; REFERENCE/DOCKET NUMBER: 0609.4070002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2540
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 152 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-460-694-4

```

```

Query Match 57.1%; Score 32; DB 2; Length 152;

```

```

Best Local Similarity 60.0%; Pred. No. 44;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMHY 10
Db 20 EEVFPPLAMNY 29

RESULT 10
US-08-460-744-4
; Sequence 4, Application US/08460744
; Patent No. 6107541
; GENERAL INFORMATION:
; APPLICANT: Arnold, Andrew
; TITLE OF INVENTION: Prad1 Cyclin and its cDNA
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,744
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McConathy, Evelyn H.
; REGISTRATION NUMBER: 35,279
; REFERENCE/DOCKET NUMBER: 0609.4070005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 152 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-460-744-4

```

```

Query Match 57.1%; Score 32; DB 3; Length 152;
Best Local Similarity 60.0%; Pred. No. 44;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 EEVVPXGMHY 10
Db 20 EEVFPPLAMNY 29

```

```

RESULT 11
US-07-667-711B-4
; Sequence 4, Application US/07667711B
; Patent No. 6110700
; GENERAL INFORMATION:
; APPLICANT: ARNOLD, ANDREW
; TITLE OF INVENTION: Prad1 Cyclin and its cDNA
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:

```


; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/667,711B
; FILING DATE: 11-MAR-1991

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MCPHAIL, DONALD R.
; REGISTRATION NUMBER: 35,811
; REFERENCE/DOCKET NUMBER: 0609.4070000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540

; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 152 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-07-667-711B-4

Query Match 57.1%; Score 32; DB 3; Length 152;
Best Local Similarity 60.0%; Pred. No. 44;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 BEVVPXGMHY 10
||| | | |
Db 20 BEVFPPLMNY 29

RESULT 12

US-08-193-977-7
; Sequence 7, Application US/08193977
; Patent No. 5625031

; GENERAL INFORMATION:
; APPLICANT: WEBSTER, KEVIN R.
; APPLICANT: COLEMAN, KEVIN G.
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF THE P33CDK2 AND
; TITLE OF INVENTION: P34CDK2 CELL CYCLE REGULATORY KINASES AND HUMAN
; TITLE OF INVENTION: PAPILLOMAVIRUS E7 ONCOPROTEIN
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301

; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/193,977
; FILING DATE: 08-FEB-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 5998-0016
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231

; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 173 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: peptide
US-08-193-977-7

Query Match 57.1%; Score 32; DB 1; Length 173;
Best Local Similarity 60.0%; Pred. No. 51;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 BEVVPXGMHY 10
||| | | |
Db 55 BEVFPPLMNY 64

RESULT 13

US-08-464-517-21
; Sequence 21, Application US/08464517
; Patent No. 5869640

; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,517
; FILING DATE:

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941

; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 189 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-464-517-21

Query Match 57.1%; Score 32; DB 2; Length 189;
Best Local Similarity 60.0%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 BEVVPXGMHY 10
||| | | |
Db 74 BEVFPPLMNY 83

RESULT 14

US-08-246-361A-21
; Sequence 21, Application US/08246361A
; Patent No. 5998582

GENERAL INFORMATION:
APPLICANT: BEACH, David H.
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/246,361A
FILING DATE: 19-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/963,308
FILING DATE: 16-OCT-1992
APPLICATION DATA:
APPLICATION NUMBER: US 07/888,178
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/701,514
FILING DATE: 16-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Matthew P. Vincent
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MII-004C
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 189 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-246-361A-21

Query Match 57.1%; Score 32; DB 2; Length 189;
Best Local Similarity 60.0%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10
Db 74 EEVFPPLAMNY 83

RESULT 15
US-08-463-772-21
Sequence 21, Application US/08463772
Patent No. 6066501
GENERAL INFORMATION:
APPLICANT: BEACH, David H.
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
*OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,772
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/963,308
FILING DATE: 16-OCT-1992
APPLICATION NUMBER: US 07/888,178
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/701,514
FILING DATE: 16-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Matthew P. Vincent
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MII-004C
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 189 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-463-772-21

Query Match 57.1%; Score 32; DB 3; Length 189;
Best Local Similarity 60.0%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10
Db 74 EEVFPPLAMNY 83

Search completed: June 4, 2003, 13:14:59
Job time : 10.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:04:09 ; Search time 28.25 Seconds
(without alignments)
51.885 Million cell updates/sec

Title: AUDET-909-3

Perfect score: 54

Sequence: 1 eevvpxgndys 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A. Geneseq 101002.*

1: /SID82/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID82/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID82/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID82/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID82/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID82/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID82/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID82/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID82/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID82/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID82/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID82/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID82/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID82/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID82/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID82/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID82/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID82/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID82/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID82/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID82/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID82/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID82/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	96.3	11	23	ABB80524
2	52	96.3	11	23	ABB80528
3	52	96.3	11	23	ABB80529
4	52	96.3	11	23	ABB80561
5	52	96.3	11	23	ABB80562
6	47	87.0	11	23	ABB80538
7	47	87.0	11	23	ABB80542
8	47	87.0	11	23	ABB80543
9	46	85.2	11	23	ABB80521
10	46	85.2	11	23	ABB80522

11	46	85.2	11	23	ABB80525	Hepatitis C virus
12	46	85.2	11	23	ABB80526	Hepatitis C virus
13	46	85.2	11	23	ABB80547	Hepatitis C virus
14	46	85.2	11	23	ABB80548	Hepatitis C virus
15	46	85.2	11	23	ABB80551	Hepatitis C virus
16	46	85.2	11	23	ABB80556	Hepatitis C virus
17	46	85.2	11	23	ABB80557	Hepatitis C virus
18	46	85.2	11	23	ABB80559	Hepatitis C virus
19	46	85.2	11	23	ABB80563	Hepatitis C virus
20	46	85.2	11	23	ABB80564	Hepatitis C virus
21	46	85.2	11	23	ABB80565	Hepatitis C virus
22	46	85.2	11	23	ABB80566	Hepatitis C virus
23	46	85.2	11	23	ABB80567	Hepatitis C virus
24	46	85.2	11	23	ABB80568	Hepatitis C virus
25	45	83.3	11	23	ABB80523	Hepatitis C virus
26	45	83.3	11	23	ABB80527	Hepatitis C virus
27	45	83.3	11	23	ABB80558	Hepatitis C virus
28	45	83.3	11	23	ABB80560	Hepatitis C virus
29	44	81.5	11	23	ABB80533	Hepatitis C virus
30	44	81.5	11	23	ABB80534	Hepatitis C virus
31	41	75.9	11	23	ABB80535	Hepatitis C virus
32	41	75.9	11	23	ABB80536	Hepatitis C virus
33	41	75.9	11	23	ABB80539	Hepatitis C virus
34	41	75.9	11	23	ABB80540	Hepatitis C virus
35	40	74.1	11	23	ABB80537	Hepatitis C virus
36	40	74.1	11	23	ABB80541	Hepatitis C virus
37	40	74.1	11	23	ABB80544	Hepatitis C virus
38	40	74.1	11	23	ABB80545	Hepatitis C virus
39	40	74.1	11	23	ABB80549	Hepatitis C virus
40	40	74.1	11	23	ABB80552	Hepatitis C virus
41	40	74.1	11	23	ABB80553	Hepatitis C virus
42	39	72.2	11	23	ABB80546	Hepatitis C virus
43	39	72.2	11	23	ABB80550	Hepatitis C virus
44	39	72.2	11	23	ABB80554	Hepatitis C virus
45	39	72.2	11	23	ABB80555	Hepatitis C virus

ALIGNMENTS

RESULT 1
ABB80524
ID ABB80524 standard; peptide; 11 AA.
XX ABB80524;
XX AC
XX DT 08-OCT-2002 (first entry)
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #4.
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide; virucide.
XX OS Synthetic.
XX Key Location/Qualifiers
FH Modified-site 1
FT /note= "N-terminal acetyl"
FT Modified-site 6
FT /note= "Norvalyl carbonyl forming keto-amide linkage with residue 7"
FT Misc-difference 9
FT /note= "D-form residue"
FT Modified-site 11
FT /note= "C-terminal amide"
XX WO200208251-A2.
PD 31-JAN-2002.
PF 19-JUL-2001; 2001WO-US23169.
XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease -
 XX
 XX Claim 17; Page 64; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 SQ Sequence 11 AA;
 Query Match 96.3%; Score 52; DB 23; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.0015;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EEVVPXGMDYS 11
 |||||
 DB 1 EEVVPXGMDYS 11
 |||||
 RESULT 2
 ABB80528
 ID ABB80528 standard; peptide; 11 AA.
 XX ABB80528;
 AC
 XX 08-OCT-2002 (first entry)
 DT
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #8.
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX Synthetic.
 OS
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8 /note= "D-form residue"
 FT Misc-difference 9 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 XX 31-JAN-2002.
 PD
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 PR
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory

PT activity useful for treating disorders associated with hepatitis C
 PT virus protease -
 XX
 XX Claim 17; Page 64; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 SQ Sequence 11 AA;
 Query Match 96.3%; Score 52; DB 23; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.0015;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EEVVPXGMDYS 11
 |||||
 DB 1 EEVVPXGMDYS 11
 |||||
 RESULT 3
 ABB80529
 ID ABB80529 standard; peptide; 11 AA.
 XX ABB80529;
 AC
 XX 08-OCT-2002 (first entry)
 DT
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #9.
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX Synthetic.
 OS
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8 /note= "D-form residue"
 FT Misc-difference 9 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 XX 31-JAN-2002.
 PD
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 PR
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease -
 XX
 XX Claim 17; Page 64; 69pp; English.
 XX The sequence represents a peptide compound of the invention having

Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EEVVPXGMDYS 11
| | | | | | | | | |
Db 1 EEVVPXGMDYS 11

RESULT 6
ABB80538
ID ABB80538 standard; peptide; 11 AA.
XX AC ABB80538;
XX DT 08-OCT-2002 (first entry)
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #18.
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX KW virucide.
XX OS Synthetic.

PH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"
FT Misc-difference 9 /note= "D-form residue"
FT Modified-site 11
FT Modified-site 11 /note= "C-terminal amide"
FT WO200208251-A2.
XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.
XX 21-JUL-2000; 2000US-220101P.
XX (CORV-) CORVAS INT INC.
XX Lim-wilby M, Levy OE, Brunck TK;
XX WPI; 2002-361643/39.
XX Novel peptide compound having hepatitis C virus protease inhibitory
FT activity useful for treating disorders associated with hepatitis C
FT virus protease -
XX Claim 17; Page 64; 69pp; English.

XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX SQ Sequence 11 AA;

Query Match 87.0%; Score 47; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.014; 1; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEVVPXGMDYS 11
| | | | | | | | | |
Db 1 EEVVPXGMDYS 11

RESULT 7
ABB80542
ID ABB80542 standard; peptide; 11 AA.
XX AC ABB80542;
XX DT 08-OCT-2002 (first entry)
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #22.
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX KW virucide.
XX OS Synthetic.

PH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"
FT Misc-difference 8 /note= "D-form residue"
FT Modified-site 11 /note= "C-terminal amide"
FT WO200208251-A2.
XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.
XX 21-JUL-2000; 2000US-220101P.
XX (CORV-) CORVAS INT INC.
XX Lim-wilby M, Levy OE, Brunck TK;
XX WPI; 2002-361643/39.
XX Novel peptide compound having hepatitis C virus protease inhibitory
FT activity useful for treating disorders associated with hepatitis C
FT virus protease -
XX Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX SQ Sequence 11 AA;

Query Match 87.0%; Score 47; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.014;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEVVPXGMDYS 11
| | | | | | | | | |
Db 1 EEVVPXGMDYS 11

RESULT 8
ABB80543
ID ABB80543 standard; peptide; 11 AA.
XX AC ABB80543;
XX DT 08-OCT-2002 (first entry)
XX

DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #23.
XX
KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"
FT
FT Misc-difference 8 /note= "D-form residue"
FT Misc-difference 9 /note= "D-form residue"
FT Modified-site 11 /note= "C-terminal amide"
XX
XX WO200208251-A2.
PN
XX 31-JAN-2002.
PD
XX 19-JUL-2001; 2001WO-US23169.
PF
XX 21-JUL-2000; 2000US-220101P.
PR
XX (CORV-) CORVAS INT INC.
PA
XX Lim-wilby M, Levy OE, Brunck TK;
PI WPI; 2002-361643/39.
XX
XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease -
XX
XX Claim 17; Page 65; 69pp; English.
PS
XX The sequence represents a peptide compound of the invention having
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX
XX Sequence 11 AA;
PS
XX Query Match 87.0%; Score 47; DB 23; Length 11;
XX Best Local Similarity 90.9%; Pred. No. 0.014;
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPXGMDYS 11
DB 1 EEVVPXGQDYS 11
DE
XX
XX RESULT 9
XX ABB80521
ID ABB80521 standard; peptide; 11 AA.
XX
XX AC ABB80521;
XX
XX 08-OCT-2002 (first entry)
DT
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #1.
DE
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX
XX Synthetic.

XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"
FT
FT Modified-site 11 /note= "C-terminal amide"
XX
XX WO200208251-A2.
PN
XX 31-JAN-2002.
PD
XX 19-JUL-2001; 2001WO-US23169.
PF
XX 21-JUL-2000; 2000US-220101P.
PR
XX (CORV-) CORVAS INT INC.
PA
XX Lim-wilby M, Levy OE, Brunck TK;
PI WPI; 2002-361643/39.
XX
XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease -
XX
XX Claim 17; Page 64; 69pp; English.
PS
XX The sequence represents a peptide compound of the invention having
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX
XX Sequence 11 AA;
PS
XX Query Match 85.2%; Score 46; DB 23; Length 11;
XX Best Local Similarity 90.9%; Pred. No. 0.022; 1; Indels 0; Gaps 0;
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPXGMDYS 11
DB 1 EEVVPXGMSYS 11
DE
XX
XX RESULT 10
XX ABB80522
ID ABB80522 standard; peptide; 11 AA.
XX
XX AC ABB80522;
XX
XX 08-OCT-2002 (first entry)
DT
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #2.
DE
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"
FT Misc-difference 9 /note= "D-form residue"
FT Modified-site 11 /note= "D-form residue"
XX
XX

FT XX /note= "C-terminal amide"
 PN WO200208251-A2.
 XX
 PD 31-JAN-2002.
 XX
 PF 19-JUL-2001; 2001WO-US23169.
 XX
 PR 21-JUL-2000; 2000US-220101P.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 PI Lim-wilby M, Levy OE, Brunck TK;
 XX
 PS WPI; 2002-361643/39.
 XX
 DR Novel peptide compound having hepatitis C virus protease inhibitory
 XX activity useful for treating disorders associated with hepatitis C
 XX virus protease -
 XX
 PS Claim 17; Page 64; 69pp; English.
 XX
 CC The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 SQ Sequence 11 AA;
 CC
 CC Query Match 85.2%; Score 46; DB 23; Length 11;
 CC Best Local Similarity 90.9%; Pred. No. 0.022;
 CC Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EVVVPXGMDYS 11
 DB ||||| ||
 1 EVVVPXGMSYS 11
 RESULT 11
 ABB80525
 ID ABB80525 standard; peptide; 11 AA.
 AC
 AC ABB80525;
 DT 08-OCT-2002 (first entry)
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #5.
 XX
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8 /note= "D-form residue"
 FT Misc-difference 9 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT
 FT WO200208251-A2.
 FT
 FT 31-JAN-2002.
 FT
 FT 19-JUL-2001; 2001WO-US23169.
 FT
 FT 21-JUL-2000; 2000US-220101P.
 FT
 FT (CORV-) CORVAS INT INC.
 FT
 FT Lim-wilby M, Levy OE, Brunck TK;
 FT
 FT WO200208251-A2.
 FT
 FT 31-JAN-2002.
 FT
 FT 19-JUL-2001; 2001WO-US23169.

PR 21-JUL-2000; 2000US-220101P.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 PI Lim-wilby M, Levy OE, Brunck TK;
 XX
 DR WPI; 2002-361643/39.
 XX
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease -
 PT
 PS Claim 17; Page 64; 69pp; English.
 XX
 CC The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 SQ Sequence 11 AA;
 CC
 CC Query Match 85.2%; Score 46; DB 23; Length 11;
 CC Best Local Similarity 90.9%; Pred. No. 0.022;
 CC Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EVVVPXGMDYS 11
 DB ||||| ||
 1 EVVVPXGMSYS 11
 RESULT 12
 ABB80526
 ID ABB80526 standard; peptide; 11 AA.
 AC
 AC ABB80526;
 DT 08-OCT-2002 (first entry)
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #6.
 XX
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 KW virucide.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8 /note= "D-form residue"
 FT Misc-difference 9 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT
 FT WO200208251-A2.
 FT
 FT 31-JAN-2002.
 FT
 FT 19-JUL-2001; 2001WO-US23169.
 FT
 FT 21-JUL-2000; 2000US-220101P.
 FT
 FT (CORV-) CORVAS INT INC.
 FT
 FT Lim-wilby M, Levy OE, Brunck TK;
 FT
 FT WO200208251-A2.
 FT
 FT 31-JAN-2002.
 FT
 FT 19-JUL-2001; 2001WO-US23169.
 FT
 FT 21-JUL-2000; 2000US-220101P.
 FT
 FT (CORV-) CORVAS INT INC.
 FT
 FT Lim-wilby M, Levy OE, Brunck TK;
 FT
 FT WO200208251-A2.
 FT
 FT 31-JAN-2002.
 FT
 FT 19-JUL-2001; 2001WO-US23169.

Query Match 85.2%; Score 46; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.022;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
| | | | | | | |
DB 1 EEVVPXGTDYS 11

RESULT 15

ABB80551
ID ABB80551 standard; peptide; 11 AA.

XX AC ABB80551;
XX DT 08-OCT-2002 (first entry)
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #31.
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX KW virucide.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT Misc-difference 9 residue 7"
FT FT /note= "D-form residue"
FT Modified-site 11
FT FT /note= "C-terminal amide"
XX WO200208251-A2.

PN 31-JAN-2002.

PD 19-JUL-2001; 2001WO-US23169.

PF 21-JUL-2000; 2000US-220101P.

PR (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

DR Novel peptide compound having hepatitis C virus protease inhibitory

PT activity useful for treating disorders associated with hepatitis C

PT virus protease

XX Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.

SQ Sequence 11 AA;

Query Match 85.2%; Score 46; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.022;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
| | | | | | | |
DB 1 EEVVPXGSDYS 11

Search completed: June 4, 2003, 13:11:13
Job time : 28.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search., using sw model

Run on: June 4, 2003, 13:07:09 ; Search time 22.5 Seconds
(without alignments)
100.734 Million cell updates/sec

Title: AUDET-909-3

Perfect score: 54

Sequence: 1 eevvpxgm dms 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertibrate.*
14: sp_unclassified.*
15: sp_rviris.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	74.1	156	3 Q12479	Q12479 saccharomyc
2	38	70.4	363	17 Q30260	Q30260 archaeoglob
3	38	70.4	1063	16 Q8RG86	Q8RG86 fusobacteri
4	36	66.7	341	10 Q22081	Q22081 citrus unsh
5	36	66.7	348	10 Q22096	Q22096 citrus unsh
6	36	66.7	452	10 Q8W568	Q8W568 arabidopsis
7	36	66.7	460	10 Q9C9T7	Q9C9T7 arabidopsis
8	36	66.7	1047	10 P93782	P93782 saccharum o
9	36	66.7	1083	10 Q9SN30	Q9SN30 arabidopsis
10	36	66.7	1084	10 Q43010	Q43010 oryza sativ
11	36	66.7	1100	10 Q8S064	Q8S064 oryza sativ
12	35	64.8	219	5 Q9GQ04	Q9GQ04 eriocheir s
13	35	64.8	253	16 Q8XPAB	Q8XPAB clostridium
14	35	64.8	298	2 Q52367	Q52367 rhizobium t
15	35	64.8	425	5 Q9XV44	Q9XV44 caenorhabdi
16	35	64.8	433	16 Q9A382	Q9A382 caulobacter

17	35	64.8	440	17 Q9YFI3	Q9YFI3 aeropyrum p
18	35	64.8	511	2 O52680	O52680 escherichia
19	35	64.8	517	16 O8XZL5	O8XZL5 ralstonia s
20	35	64.8	745	5 Q9SP46	Q9SP46 carinus ma
21	35	64.8	1031	5 Q9UA83	Q9UA83 callinectes
22	35	64.8	1150	5 O17704	O17704 caenorhabdi
23	35	64.8	1410	2 O52673	O52673 escherichia
24	35	64.8	1420	2 O52666	O52666 escherichia
25	35	64.8	1474	17 O27146	O27146 methanobact
26	35	64.8	1828	16 Q98K29	Q98K29 rhizobium l
27	35	64.8	2778	5 Q9V9T6	Q9V9T6 drosophila
28	34.5	63.9	748	4 Q8TBJ7	Q8TBJ7 homo sapien
29	34	63.0	143	5 Q9VSY8	Q9VSY8 drosophila
30	34	63.0	154	10 Q9SBB8	Q9SBB8 oryza sativ
31	34	63.0	215	16 Q8R9L5	Q8R9L5 thermaanaer
32	34	63.0	290	16 Q8U7J0	Q8U7J0 agrobacteri
33	34	63.0	296	17 Q9YET8	Q9YET8 aeropyrum p
34	34	63.0	357	17 O29920	O29920 archaeoglob
35	34	63.0	366	17 Q94511	Q94511 archaeoglob
36	34	63.0	387	16 Q98FX1	Q98FX1 rhizobium l
37	34	63.0	543	3 Q8TFP4	Q8TFP4 trichoderma
38	34	63.0	558	16 Q8R8Z2	Q8R8Z2 thermaanaer
39	34	63.0	565	16 Q9CIN1	Q9CIN1 lactococcus
40	34	63.0	587	16 Q9JZP8	Q9JZP8 neisseria m
41	34	63.0	906	10 Q9ZSY4	Q9ZSY4 arabidopsis
42	34	63.0	908	10 Q9FJK8	Q9FJK8 arabidopsis
43	34	63.0	908	10 Q8W4J9	Q8W4J9 arabidopsis
44	34	63.0	908	10 Q9ZSY3	Q9ZSY3 arabidopsis
45	34	63.0	909	10 Q9WSA1	Q9WSA1 arabidopsis

ALIGNMENTS

RESULT 1

Q12479 PRELIMINARY; PRT; 156 AA.
AC Q12479;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE ORF YOR013W.
GN YOR013W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA De haan M., Grivell L.A., Maarse A.C.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MIPS;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=FY1679;
RA De haan M., Maarse A.C., Grivell L.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=FY1679;
RX MEDLINE=94019318; PubMed=8413243;
RA Dumont M.E., Schlichter J.B., Cardillo T.S., Hayes M.K., Bethlenny G.,
RT "CYC encodes a factor involved in mitochondrial import of yeast
cytochrome c.";
RL Mol. Cell. Biol. 13:6442-6451(1993).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=FY1679;
RX MEDLINE=94169519; PubMed=7764548;

RA Lee Y.S., Shimizu J., Yoda K., Yamasaki M.;
 RT "Molecular cloning of a gene, DHS1, which complements a drug-
 sensitive mutation of the yeast *Saccharomyces cerevisiae*.";
 RL Biosci. Biotechnol. Biochem. 58:391-395(1994).
 DR EMBL; Z74920; CAA99201.1; -;
 DR EMBL; X87331; CAA60762.1; -;
 DR SGD; S0005539; YOR013W.
 SQ SEQUENCE 156 AA; 17881 MW; 380442B74C272B41 CRC64;

Query Match 74.1%; Score 40; DB 3; Length 156;
 Best Local Similarity 77.8%; Pred. No. 1.7;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMDY 10
 Db 50 EVNPLGMDY 58

RESULT 2

O30260 PRELIMINARY; PRT; 363 AA.
 ID O30260
 AC O30260
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical protein AF2411.
 GN AF2411.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 OX NCBI_TaxID=2234;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 reducing archaeon *Archaeoglobus fulgidus*.";
 RL Nature 390:364-370(1997).
 DR EMBL; AE001109; AAB91255.1; -;
 DR TIGR; AF2411; -;
 DR InterPro; IPR002103; Bac_luciferase.
 DR Pfam; PF00296; bac_luciferase; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 363 AA; 41736 MW; 0E976EAF788F4803 CRC64;

Query Match 70.4%; Score 38; DB 17; Length 363;
 Best Local Similarity 54.8%; Pred. No. 12;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVVPXGMDY 11
 Db 120 ENIVPYGIDFS 130

RESULT 3

O8RG86 PRELIMINARY; PRT; 1063 AA.
 ID O8RG86
 AC O8RG86
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Carbanoyl-phosphate synthase large chain (EC 6.3.5.5).

GN FN0422.
 OS Fusobacterium nucleatum (subsp. nucleatum).
 OC Bacteria; Fusobacteria; Fusobacterium.
 OX NCBI_TaxID=76856;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 25586;
 RX MEDLINE=21886394; PubMed=11889109;
 RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
 RA Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
 RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
 RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
 RA Fongstein M., Kyrpides N., Overbeek R.;
 RT "Genome sequence and analysis of the oral bacterium *Fusobacterium*
 RT nucleatum strain ATCC 25586.";
 RL J. Bacteriol. 184:2005-2018(2002).
 RL EMBL; AE010554; AAL94625.1; -;
 KW Ligase; Complete proteome.
 SQ SEQUENCE 1063 AA; 118008 MW; 39700E10B7CCE411 CRC64;

Query Match 70.4%; Score 38; DB 16; Length 1063;
 Best Local Similarity 60.0%; Pred. No. 39;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMDY 11
 Db 195 EIVPGLNYS 204

RESULT 4

O22081 PRELIMINARY; PRT; 341 AA.
 ID O22081
 AC O22081
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Sucrose-phosphate synthase (Fragment).
 GN CITSPS2.
 OS Citrus unshiu (Satsuma orange).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Sapindales; Rutaceae; Citrus.
 OX NCBI_TaxID=55188;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. MIYAGAWA-WASE; TISSUE=JUICE SACS AND SEGMENT EPIDERMIS;
 RX MEDLINE=96439842; PubMed=8842155;
 RA Konatsu A., Takanokura Y., Omura M., Akihama T.;
 RT "Cloning and molecular analysis of cDNA encoding three sucrose
 RT phosphate synthase isoforms from a citrus fruit (*Citrus unshiu*
 RT Marc.)."
 RL Mol. Gen. Genet. 252:346-351(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. MIYAGAWA-WASE; TISSUE=JUICE SACS AND SEGMENT EPIDERMIS;
 RA Konatsu A., Takanokura Y., Moriguchi T., Omura M., Akihama T.;
 RT "Differential expression of three sucrose-phosphate synthase isoforms
 RT during sucrose accumulation in citrus fruits (*Citrus unshiu* Marc.)."
 RL Plant Sci. 140:169-178(1999).
 DR EMBL; AB006319; BAA23215.1; -;
 FT NON TER 1
 FT NON TER 341
 SQ SEQUENCE 341 AA; 38136 MW; 61417A69C4560777 CRC64;

Query Match 66.7%; Score 36; DB 10; Length 341;
 Best Local Similarity 66.7%; Pred. No. 29;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDY 11
 Db 228 VVPGMDYFS 236

```

RESULT 5
OZ2096
ID OZ2096 PRELIMINARY; PRT; 348 AA.
AC OZ2096;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Sucrose-phosphate synthase (Fragment).
OS CITRUS.
OS Citrus unshiu (Satsuma orange).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Sapindales; Rutaceae; Citrus.
OX NCBI_TaxID=55188;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=JUICE SACS AND SEGMENT EPIDERMIS;
RX MEDLINE=96439842; PubMed=8842155;
RA Komatsu A., Takanokura Y., Omura M., Akihama T.;
RT "Cloning and molecular analysis of cDNA encoding three sucrose
RT phosphate synthase isoforms from a citrus fruit (Citrus unshiu
RT Marc.).";
RN Mol. Gen. Genet. 252:346-351(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=JUICE SACS AND SEGMENT EPIDERMIS;
RA Komatsu A., Takanokura Y., Moriguchi T., Omura M., Akihama T.;
RT "Differential expression of three sucrose-phosphate synthase isoforms
RT during sucrose accumulation in citrus fruits (Citrus unshiu Marc.).";
RN Plant Sci. 140:169-178(1999).
DR EMBL; AB006660; BAA22071.1; -.
FT NON TER 1
FT NON TER 348
SQ SEQUENCE 348 AA; 38556 MW; EE1C21EBA6FF5C5E CRC64;

Query Match 66.7%; Score 36; DB 10; Length 348;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
DB 234 VVPGMDPS 242

RESULT 6
Q8W568 PRELIMINARY; PRT; 452 AA.
AC Q8W568;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE At1g73750/F25P22.17.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF419606; AAL31938.1; -.
DR InterPro; IPR002471; Prol_endopep_ser.
DR InterPro; IPR000379; Ser_estrs_site.
DR PROSITE; PS00708; PRO_ENDOPEP_SER; UNKNOWN_1.

SQ SEQUENCE 452 AA; 49682 MW; A1S9955B21742C4A CRC64;

Query Match 66.7%; Score 36; DB 10; Length 452;
Best Local Similarity 70.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
DB 210 EEDVPSAMDY 219

RESULT 7
Q9C9T7 PRELIMINARY; PRT; 460 AA.
AC Q9C9T7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical 50.6 kDa protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:816-820(2000).
DR EMBL; AC012679; AAG52073.1; -.
DR InterPro; IPR002471; Prol_endopep_ser.
DR InterPro; IPR000379; Ser_estrs_site.
DR PROSITE; PS00708; PRO_ENDOPEP_SER; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 460 AA; 50564 MW; E94B27B5C4B249EC CRC64;

Query Match 66.7%; Score 36; DB 10; Length 460;
Best Local Similarity 70.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
DB 218 EEDVPSAMDY 227

RESULT 8
P93782 PRELIMINARY; PRT; 1047 AA.
AC P93782;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Sucrose-phosphate synthase (EC 2.4.1.14) (Fragment).
OX NCBI_TaxID=55188;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=JUICE SACS AND SEGMENT EPIDERMIS;
RX MEDLINE=96439842; PubMed=8842155;
RA Komatsu A., Takanokura Y., Omura M., Akihama T.;
RT "Cloning and molecular analysis of cDNA encoding three sucrose
RT phosphate synthase isoforms from a citrus fruit (Citrus unshiu
RT Marc.).";
RN Mol. Gen. Genet. 252:346-351(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=JUICE SACS AND SEGMENT EPIDERMIS;
RA Komatsu A., Takanokura Y., Moriguchi T., Omura M., Akihama T.;
RT "Differential expression of three sucrose-phosphate synthase isoforms
RT during sucrose accumulation in citrus fruits (Citrus unshiu Marc.).";
RN Plant Sci. 140:169-178(1999).
DR EMBL; AB006660; BAA22071.1; -.
FT NON TER 1
FT NON TER 348
SQ SEQUENCE 348 AA; 38556 MW; EE1C21EBA6FF5C5E CRC64;

Query Match 66.7%; Score 36; DB 10; Length 348;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
DB 234 VVPGMDPS 242

```

```

OS Saccharum officinarum (Sugarcane).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoidae; Andropogoneae; Saccharum.
OX NCBI_TaxID=4547;
RN [1]_TaxID=4547;
RP SEQUENCE FROM N.A.
RC TISSUE=LEAF;
RA Sugiharto B., Sakakibara H., Sugiyama T.;
RT "Differential Expression of Two Genes for Sucrose-Phosphate Synthase
RT in Sugarcane: Molecular Cloning of the cDNAs and Comparative Analysis
RT of Gene Expression."
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB001337; BAA19241.1; -.
DR InterPro; IPR001296; Glycos transf 1.
DR Pfam; PF00534; Glycos transf 1; 1.
KW Glycosyltransferase; Transferase.
FT NON_TER
SQ SEQUENCE 1047 AA; 116379 MW; D0EDB34961E1D83D CRC64;

Query Match 66.7%; Score 36; DB 10; Length 1047;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
Db 414 VPPGMDFS 422

RESULT 9
Q9SN30 PRELIMINARY; PRT; 1083 AA.
AC Q9SN30;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Sucrose-phosphate synthase-like protein (EC 2.4.1.14).
GN F28M11.40 OR AT4G10120.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,
RA Mayer K.F.X., Lemcke K., Schueller C.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL049487; CAB39764.1; -.
DR EMBL; AL161516; CAB78135.1; -.
DR InterPro; IPR001296; Glycos transf 1.
DR Pfam; PF00534; Glycos transf 1; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 1083 AA; 122688 MW; EFC126FCA2137BB1 CRC64;

Query Match 66.7%; Score 36; DB 10; Length 1083;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
Db 483 VPPGMDFS 491

RESULT 10

```

```

Q43010 PRELIMINARY; PRT; 1084 AA.
AC Q43010;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE L-lactate dehydrogenase (EC 1.1.1.27) (LDH).
GN SPS1.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, JAPONICA;
RA Sakamoto M., Satozawa T., Kishimoto N., Higo K., Shimada H.,
RA Fujimura T.;
RT "Structure and RFLP mapping of a rice sucrose phosphate synthase (SPS)
RT gene that is specifically expressed in the source organ."
RL Plant Sci. 112:207-217 (1995).
CC -|- CATALYTIC ACTIVITY: (S)-LACTATE + NAD(+) = PYRUVATE + NADH.
CC -|- PATHWAY: ANAEROBIC GLYCOLYSIS; FINAL STEP.
CC -|- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
DR EMBL; D45890; BAA08304.1; -.
DR InterPro; IPR001296; Glycos transf 1.
DR Pfam; PF00534; Glycos transf 1; 1.
DR PROSITE; PS00064; LDH; 1.
KW GLYCOLYSIS; NAD; Oxidoreductase.
SQ SEQUENCE 1084 AA; 119417 MW; 4D0D79AFEDD2F92B CRC64;

Query Match 66.7%; Score 36; DB 10; Length 1084;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
Db 453 VPPGMDFS 461

RESULT 11
Q8S064 PRELIMINARY; PRT; 1100 AA.
AC Q8S064;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative sucrose-phosphate synthase.
GN P0678F11.14.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
RT clone: P0678F11."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003437; BAB86107.1; -.
SQ SEQUENCE 1100 AA; 121170 MW; 084F4604BA389CAD CRC64;

Query Match 66.7%; Score 36; DB 10; Length 1100;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
Db 469 VPPGMDFS 477

```

RESULT 12

Q9GQ04 PRELIMINARY; PRT; 219 AA.
 AC Q9GQ04, 2001 (TREMELrel. 16, Created)
 DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
 DE Na+/K+/2Cl-cotransporter (Fragment).
 OS Eriocheir sinensis (Chinese mitten crab).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
 OC Eubrachyura; Grapsoidae; Varunidae; Eriocheir.
 OX NCBI_TaxID=95602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=GILL;
 RA Wehrhach D., Towle D.W.;
 RT "Na+/H+-exchanger and Na+/K+/2Cl- cotransporter are expressed in
 RT gills of the euryhaline Chinese crab *Eriocheir sinensis*.";
 RL Comp. Biochem. Physiol. 126:S158-S158(2000).
 DR EMBL: AF301160; AAC39938.1; -.
 FT NON_TER 1
 FT NON_TER 1
 FT NON_TER 219
 SQ SEQUENCE 219 AA; 24159 MW; 599442DA26F0D3F1 CRC64;

Query Match 64.8%; Score 35; DB 5; Length 219;
 Best Local Similarity 75.0%; Pred. No. 29;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 VPXGMDYS 11

Db 107 VPQGLDYS 114

RESULT 13

Q8XPAB PRELIMINARY; PRT; 253 AA.
 AC Q8XPAB;
 DT 01-MAR-2002 (TREMELrel. 20, Created)
 DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)
 DE Hypothetical protein CPE0057.
 GN CPE0057.
 OS Clostridium perfringens.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Clostridiales; Clostridiaceae; Clostridium.
 OX NCBI_TaxID=1502;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=13 / TYPE A;
 RX PubMed=11792842;
 RA Shimizu T., Ohtani K., Hirakawa H., Ohehima K., Yamashita A.,
 RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
 RT "Complete genome sequence of Clostridium perfringens, an anaerobic
 RT flesh-eater.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
 DR EMBL: AP003185; BAB79763.1; -.
 DR InterPro; IPR000205; NAD binding.
 DR InterPro; IPR000594; Thif_domain.
 DR Pfam; PF00899; Thif; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 253 AA; 27714 MW; 814DF79D3E0D7486 CRC64;

Query Match 64.8%; Score 35; DB 16; Length 253;
 Best Local Similarity 50.0%; Pred. No. 34;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 BEVFXGMDY 10

Db 108 BEIIPDDVDY 117

RESULT 14

Q52367 PRELIMINARY; PRT; 298 AA.
 AC Q52367;
 DT 01-JUN-1998 (TREMELrel. 06, Created)
 DT 01-JUN-1998 (TREMELrel. 06, Last sequence update)
 DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
 DE Aryl-alcohol dehydrogenase homolog (Fragment).
 GN XYLB1.
 OS Rhizobium tropici.
 OG Plasmid pRtrCFN299a.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=398;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CFN299;
 RA Rosenbluth M., Hynes M.F., Martinez-Romero E.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- COPACTOR: ZINC (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE
 CC FAMILY.
 DR EMBL: AF036920; AAC04779.1; -.
 DR HSP; P07846; 1SDG.
 DR InterPro; IPR002328; ADH_zinc.
 DR InterPro; IPR002085; ADH_zn_family.
 DR InterPro; IPR000205; NAD_binding.
 DR Pfam; PF00107; adh_zinc; 1.
 DR PROSITE; PS00059; ADH_ZINC; 1.
 KW Oxidoreductase; Plasmid; Zinc.
 FT NON_TER 298
 FT NON_TER 298
 SQ SEQUENCE 298 AA; 31092 MW; 49B2F0117C33AE87 CRC64;

Query Match 64.8%; Score 35; DB 2; Length 298;
 Best Local Similarity 50.0%; Pred. No. 41;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 EVVFXGMDYS 11

Db 250 EIIPEGADFS 259

RESULT 15

Q9XVK4 PRELIMINARY; PRT; 425 AA.
 AC Q9XVK4;
 DT 01-NOV-1999 (TREMELrel. 12, Created)
 DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)
 DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
 DE R10D12.10 protein.
 GN R10D12.10.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Percy C.M.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode *C. elegans*: A platform for
 RT investigating biology.";
 RL Science 282:2012-2018(1998).
 DR EMBL: Z81109; CAB03241.1; -.
 DR InterPro; IPR000719; Euk_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Transferase.
 SQ SEQUENCE 425 AA; 49410 MW; 5D96E29B08C8E9D6 CRC64;

Query Match 64.8%; Score 35; DB 5; Length 425;
Best Local Similarity 50.0%; Pred. No. 61;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 BEVVPXGMDY 10
|::|||:
Db 335 EQIVPGGLQY 344

Search completed: June 4, 2003, 13:13:26
Job time : 23.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:04:34 ; Search time 6.25 seconds
(without alignments)
72.998 Million cell updates/sec

Title: AUDET-909-3

Perfect score: 54

Sequence: 1 eevvpxgmids 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	38	70.4	1058	1 CARB_FUSNN	Q8rg86 fusobacteri
2	36	66.7	102	1 Y1LK_TYDVA	P31619 tobacco yel
3	36	66.7	1049	1 SPS_ORISA	Q43802 oryza sativ
4	36	66.7	1068	1 SPS_WAIZE	P31927 zea mays (m
5	36	66.7	1081	1 SPS2_CRAPL	O04933 cratostig
6	35	64.8	2747	1 EAF_DROME	P55824 drosophila
7	34.5	63.9	748	1 KHLI_HUMAN	Q8nr64 homo sapien
8	34	63.0	154	1 PLAS_ORISA	P20423 oryza sativ
9	34	63.0	155	1 PLAS_HORVU	P08248 hordeum vul
10	34	63.0	168	1 PLAT_POPNI	P11970 populus nig
11	34	63.0	566	1 SYPB_BORBU	P94283 borrelia bu
12	33	61.1	276	1 Y939_METUA	Q58349 methanococ
13	33	61.1	283	1 PANC_PSEAE	Q9hv69 pseudomonas
14	33	61.1	394	1 HMPA_VIBCH	Q9kmy3 vibrio chol
15	33	61.1	421	1 ACDM_RAT	P08503 rattus norv
16	33	61.1	421	1 ECB2_HALEL	O52250 halomonas e
17	33	61.1	423	1 ECB1_HALEL	Q9zeu7 halomonas e
18	33	61.1	787	1 ECB2_HUMAN	G60344 homo sapien
19	33	61.1	801	1 FGR3_MOUSE	Q61851 mus musculu
20	33	61.1	806	1 CEK2_CHICK	P18460 gallus gall
21	33	61.1	877	1 SULH_SCHPO	O74377 schizosacch
22	33	61.1	982	1 ENV_SVB3L	P27399 smian foam
23	33	61.1	1401	1 RPOC_VIBCH	Q9kv29 vibrio chol
24	33	61.1	2717	1 ZEP1_HUMAN	P15822 homo sapien
25	32.5	60.2	472	1 ET2A_XENLA	P19102 xenopus lae
26	32	59.3	97	1 PLAS_DAUCA	P20422 gaucous caro
27	32	59.3	175	1 HES3_RAT	Q04667 rattus norv
28	32	59.3	231	1 ARAD_ECOLI	P08203 escherichia
29	32	59.3	231	1 ARAD_SALTY	P06190 salmonella
30	32	59.3	233	1 HIS9_THEMEA	Q9wzr1 thermotoga
31	32	59.3	288	1 CGD2_RAT	Q04827 rattus norv
32	32	59.3	289	1 CGD2_HUMAN	P30279 homo sapien
33	32	59.3	289	1 CGD2_MOUSE	P30280 mus musculu

RESULT 1					
CARB_FUSNN	STANDARD;	PRT; 1058 AA.			
AC	Q8RG86;				
DT	15-JUN-2002 (Rel. 41, Created)				
DT	15-JUN-2002 (Rel. 41, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-phosphate synthetase ammonia chain).				
GN	CARB OR FN0422.				
OS	Fusobacterium nucleatum (subsp. nucleatum).				
OC	Bacteria; Fusobacteria; Fusobacterium.				
OX	NCBI_TaxID=76856;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=ATCC 25586;				
RX	MEDLINE=21886394; PubMed=11889109;				
RA	Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A., Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L., Vasileva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A., Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R., Fonstein M., Kyriides N., Overbeek R.;				
RA	"Genome sequence and analysis of the oral bacterium Fusobacterium nucleatum strain ATCC 25586."				
RT	J. Bacteriol. 184:2005-2018(2002).				
RL	-I- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP + phosphate + L-glutamate + carbamoyl phosphate.				
CC	-I- COFACTOR: Binds three manganese ions (By similarity).				
CC	-I- PATHWAY: Arginine biosynthesis.				
CC	-I- PATHWAY: Pyrimidine biosynthesis; first step.				
CC	-I- SUBUNIT: Composed of two chains; the small (or glutamine) chain promotes the hydrolysis of glutamine to ammonia, which is used by the large (or ammonia) chain to synthesize carbamoyl phosphate (By similarity).				
CC	-I- SIMILARITY: BELONGS TO THE CARB FAMILY.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	EMBL; A010554; AAL94625.1; ALT INIT.				
DR	InterPro; IPR005483; CPase_L.				
DR	InterPro; IPR005479; CPase_L_D2.				
DR	InterPro; IPR005480; CPase_L_D3.				
DR	InterPro; IPR005481; CPase_L_N.				
DR	InterPro; IPR004362; MGS_like.				
DR	Pfam; PF00289; CPase_L_Chain; 2.				
DR	Pfam; PF02786; CPase_L_D2; 2.				
DR	Pfam; PF02787; CPase_L_D3; 1.				
DR	Pfam; PF02142; MGS; 1.				
DR	PRINTS; PR00098; CPASE.				
DR	PROSITE; PS00866; CPASE_1; 2.				

Q90459 brachydanio
P50755 xenopus lae
P49706 gallus gall
P53782 xenopus lae
P55169 gallus gall
P30281 homo sapien
P24385 homo sapien
P25322 mus musculu
P39948 rattus norv
O30640 methanosarc
Q91712 xenopus lae
Q91910 xenopus lae

ALIGNMENTS

```

DR PROSITE; PS00867; CFSASE.2; 2.
KW Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;
KW ATP-binding; Manganese; Complete proteome.
FT DOMAIN 1 401 CARBOXYPHOSPHATE SYNTHETIC DOMAIN.
FT DOMAIN 402 546 OLIGOMERIZATION DOMAIN.
FT DOMAIN 547 929 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.
FT DOMAIN 930 1058 ALLOSTERIC DOMAIN.
FT REPEAT 1 546
FT REPEAT 547 1058
FT NP_BIND 153 210 ATP (POTENTIAL).
FT NP_BIND 302 352 ATP (POTENTIAL).
FT METAL 284 288 MANGANESE 1 (BY SIMILARITY).
FT METAL 298 298 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 300 300 MANGANESE 2 (BY SIMILARITY).
FT METAL 820 820 MANGANESE 3 (BY SIMILARITY).
FT METAL 832 832 MANGANESE 3 (BY SIMILARITY).
SQ SEQUENCE 1058 AA; 117451 MW; ED7037AF77C1E339F CRC64;

Query Match 70.4%; Score 38; DB 1; Length 1058;
Best Local Similarity 60.0%; Pred. No. 7.8;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPGXGMDYS 11
Db 190 EIVPGLNYS 199

RESULT 2
Y1LK TYDVA
ID Y1LK TYDVA STANDARD; PRT; 102 AA.
AC P31619;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-OCT-1993 (Rel. 27, Last annotation update)
DE Hypothetical 11.2 kDa protein.
GN Vi
OS Tobacco yellow dwarf virus (strain Australia) (TYDV).
OC Viruses; ssDNA viruses; Geminiviridae; Mastrevirus.
OX NCBI_TaxID=31599;
RN [1]
RX MEDLINE=92188538; PubMed=1546458;
RA Morris B.A.M., Richardson K.A., Haley A., Zhan X., Thomas J.E.;
RT "The nucleotide sequence of the infectious cloned DNA component of
RT tobacco yellow dwarf virus reveals features of geminiviruses
RT infecting monocotyledonous plants.";
RL Virology 187:633-642(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M81103; AAA47947.1; -.
DR PIR; A42452;
DR InterPro; IPR002621; Gemini_mov.
DR Pfam; PF01708; Gemini_mov; 1.
KW Hypothetical protein.
SQ SEQUENCE 102 AA; 11178 MW; A40ECF1E0AF55B67 CRC64;

Query Match 66.7%; Score 36; DB 1; Length 102;
Best Local Similarity 60.0%; Pred. No. 1.7;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPGXGMDYS 11
Db 7 QVVPFGINS 16

RESULT 3

```

```

SPS ORYSA
ID SPS ORYSA STANDARD; PRT; 1049 AA.
AC Q43802;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Sucrose-phosphate synthase (EC 2.4.1.14)
DE (UDP-glucose-fructose-phosphate glucosyltransferase).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Erihartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Indica-IR36; TISSUE=Leaf;
RX MEDLINE=96235138; PubMed=8666248;
RA Valdez-Alarcon J.J., Ferrando M., Jimenez-Moraila B.,
RA Herrera-Estralla L.;
RT "Characterization of a rice sucrose-phosphate synthase-encoding
RT gene.";
RL Gene 170:217-222(1996).
CC -----
CC -!- FUNCTION: INVOLVED IN THE REGULATION OF CARBON PARTITIONING IN
CC THE LEAVES OF PLANTS. MAY REGULATE THE SYNTHESIS OF SUCROSE AND
CC THEREFORE PLAY A MAJOR ROLE AS A LIMITING FACTOR IN THE EXPORT OF
CC PHOTOASSIMILATES OUT OF THE LEAF.
CC -!- CATALYTIC ACTIVITY: UDP-glucose + D-fructose 6-phosphate = UDP +
CC sucrose 6-phosphate.
CC -!- ENZYME REGULATION: ACTIVITY REGULATED BY PHOSPHORYLATION AND
CC MODERATED BY CONCENTRATION OF METABOLITES AND LIGHT.
CC -!- PATHWAY: Sucrose synthesis.
CC -!- SUBUNIT: HOMODIMER OR HOMOTETRAMER (BY SIMILARITY).
CC -!- PTM: PHOSPHORYLATED. HOWEVER, PHOSPHORYLATION IS NOT ESSENTIAL FOR
CC ENZYME FUNCTION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U33175; AAC49379.1; -.
DR InterPro; IPR001296; Glycos_transf_1.
DR Pfam; PF00534; Glycos_transf_1; 1.
KW Transferase; Glycosyltransferase; Phosphorylation.
FT DOMAIN 22 29 POLY-GLY.
FT DOMAIN 695 698 POLY-GLU.
FT DOMAIN 775 779 POLY-ARG.
SQ SEQUENCE 1049 AA; 116455 MW; ED862E2819AA4B04 CRC64;

Query Match 66.7%; Score 36; DB 1; Length 1049;
Best Local Similarity 66.7%; Pred. No. 20;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
Db 436 VIPPGMDFS 444

RESULT 4
SPS MAIZE
ID SPS MAIZE STANDARD; PRT; 1068 AA.
AC P31927;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Sucrose-phosphate synthase (EC 2.4.1.14) (UDP-glucose-fructose-
DE phosphate glucosyltransferase).
GN SPS
OS Zea mays (Maize).

```

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 71-74; 206-212; 471-481 AND
 RP 872-892.
 RC STRAIN=CV. PIONEER 3184; TISSUE=Leaf;
 RX MEDLINE=92338837; PubMed=1840396;
 RA Worrell A.C., Bruneau J.-M., Summerfelt K., Boersig M., Voelker T.A.;
 RA "Expression of a maize sucrose phosphate synthase in tomato alters
 RT leaf carbohydrate partitioning."
 RL Plant Cell 3:1121-1130(1991).
 CC -!- FUNCTION: INVOLVED IN THE REGULATION OF CARBON PARTITIONING IN
 CC THE LEAVES OF PLANTS. MAY REGULATE THE SYNTHESIS OF SUCROSE AND
 CC THEREFORE PLAY A MAJOR ROLE AS A LIMITING FACTOR IN THE EXPORT OF
 CC PHOTOASSIMILATES OUT OF THE LEAF.
 CC -!- CATALYTIC ACTIVITY: UDP-glucose + D-fructose 6-phosphate = UDP +
 CC sucrose 6-phosphate.
 CC -!- ENZYME REGULATION: ACTIVITY REGULATED BY PHOSPHORYLATION AND
 CC MODERATED BY CONCENTRATION OF METABOLITES AND LIGHT.
 CC -!- PATHWAY: Sucrose synthesis.
 CC -!- SUBUNIT: HOMODIMER OR HOMOTETRAMER.
 CC -!- PTM: PHOSPHORYLATED. HOWEVER, PHOSPHORYLATION IS NOT ESSENTIAL FOR
 CC ENZYME FUNCTION.
 CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 1.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M97550; AAA33513.1; -;
 DR PIR; JQ1329; JQ1329.
 DR MaizeDB; 25294; -;
 DR InterPro; IPR001296; Glycosyltransf_1; 1.
 DR Pfam; PF00534; Glycosyltransf_1; 1.
 KW Transferase; Glycosyltransferase; Phosphorylation.
 FT DOMAIN 25 31 POLY-GLY.
 FT SEQUENCE 1068 AA; 118575 MW; 074679B59A1D282 CRC64;
 SQ
 Query Match 66.7%; Score 36; DB 1; Length 1068;
 Best Local Similarity 66.7%; Pred. No. 21;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 3 VVPXGMDYS 11
 DB 435 VIPPGMDFS 443
 ||:|||||
 RESULT 5
 SP22 CRAPL STANDARD; PRT; 1081 AA.
 AC 004933;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Sucrose-phosphate synthase 2 (EC 2.4.1.14) (UDP-glucose-fructose-
 DE phosphate glucosyltransferase 2).
 GN SP2.
 OS Craterostigma plantaginum.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Lamiales; Lamiales incertae sedis; Torneaeae;
 OC Craterostigma.
 OX NCBI_TaxID=4153;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97451773; PubMed=9306694;

RA Ingram J., Chandler J.W., Gallagher L., Salami F., Bartels D.;
 RT "Analysis of cDNA clones encoding sucrose-phosphate synthase in
 RT relation to sugar interconversions associated with dehydration in the
 RT resurrection plant *Craterostigma plantaginum* Hochst.";
 RL Plant Physiol 115:113-121(1997).
 CC -!- FUNCTION: INVOLVED IN THE REGULATION OF CARBON PARTITIONING IN
 CC THE LEAVES OF PLANTS. MAY REGULATE THE SYNTHESIS OF SUCROSE AND
 CC THEREFORE PLAY A MAJOR ROLE AS A LIMITING FACTOR IN THE EXPORT OF
 CC PHOTOASSIMILATES OUT OF THE LEAF.
 CC -!- CATALYTIC ACTIVITY: UDP-glucose + D-fructose 6-phosphate = UDP +
 CC sucrose 6-phosphate.
 CC -!- ENZYME REGULATION: ACTIVITY REGULATED BY PHOSPHORYLATION AND
 CC MODERATED BY CONCENTRATION OF METABOLITES AND LIGHT.
 CC -!- PATHWAY: Sucrose synthesis.
 CC -!- SUBUNIT: HOMODIMER OR HOMOTETRAMER (BY SIMILARITY).
 CC -!- PTM: PHOSPHORYLATED. HOWEVER, PHOSPHORYLATION IS NOT ESSENTIAL FOR
 CC ENZYME FUNCTION (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 1.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; Y11795; CAA72491.1; -;
 DR InterPro; IPR001296; Glycosyltransf_1; 1.
 DR Pfam; PF00534; Glycosyltransf_1; 1.
 KW Transferase; Glycosyltransferase; Phosphorylation; Multigene family.
 FT DOMAIN 245 248 POLY-SER.
 FT DOMAIN 255 264 POLY-GLU.
 FT DOMAIN 787 790 POLY-ARG.
 FT SEQUENCE 1081 AA; 120933 MW; DD142DC2F1A72900 CRC64;
 SQ
 Query Match 66.7%; Score 36; DB 1; Length 1081;
 Best Local Similarity 66.7%; Pred. No. 21;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 3 VVPXGMDYS 11
 DB 445 VIPPGMDFS 453
 ||:|||||
 RESULT 6
 FAF_DROME STANDARD; PRT; 2747 AA.
 AC P55824;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable ubiquitin carboxyl-terminal hydrolase FAF (EC 3.1.2.15)
 DE (Ubiquitin thiolesterase FAF) (Ubiquitin-specific processing protease
 DE FAF) (Deubiquitinating enzyme FAF) (Fat facets protein).
 GN FAF.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93202020; PubMed=1295747;
 RA Fischer-Vize J.A., Rubin G.M., Lehmann R.;
 RT "The fat facets gene is required for Drosophila eye and embryo
 RL development."
 RL Development 116:985-1000(1992).
 CC -!- FUNCTION: REQUIRED FOR EYE AND EMBRYO DEVELOPMENT, AND PLAYS A
 CC ROLE IN COMPOUND EYE ASSEMBLY AND OCULEGENESIS RESPECTIVELY. IN THE
 CC LARVAL EYE DISKS, CELLS OUTSIDE THE ASSEMBLING FACETS REQUIRE THIS
 CC PROTEIN FOR SHORT-RANGE CELL INTERACTIONS THAT PREVENT THE MYSTERY
 CC CELLS FROM BECOMING PHOTORECEPTORS. IT IS ALSO REQUIRED FOR

RT "Molecular cloning and characterization of plastocyanin precursor in rice.";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 58-154.
 RC STRAIN=cv. Japonica;
 RX MEDLINE=8938623; PubMed=2780537;
 RA Yano H., Kamo M., Tsugita A., Aso K., Nozu Y.;
 RT "The amino acid sequence of plastocyanin from rice (*Oryza sativa*,
 RT subspecies japonica).";
 RL Protein Seq. Data Anal. 2:385-389(1989).
 CC -!- FUNCTION: Participates in electron transfer between P700 and the
 CC cytochrome b6-f complex in photosystem I.
 CC -!- SUBCELLULAR LOCATION: LOOSELY BOUND TO THE INNER THYLAKOID
 CC MEMBRANE SURFACE IN CHLOROPLASTS.
 CC -!- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF093636; AAC78108.1; -;
 DR PIR; S06105; S06105.
 DR JTO352; JTO352.
 DR HSP; P00289; 2PCF.
 DR InterPro; IPR000923; BlueCu_1.
 DR InterPro; IPR001235; Copper_blue.
 DR Pfam; PF00127; copper-bind; 1.
 DR PRINTS; PR00156; COPPERBLUE.
 DR ProDom; PD001235; Copper blue; 1.
 DR PROSITE; PS00196; COPPER_BLUE; 1.
 KW Chloroplast; Electron transport; Copper; Thylakoid; Membrane;
 KW Transit peptide.
 FT TRANSIT 1 57 CHLOROPLAST.
 FT CHAIN 58 154 PLASTOCYANIN.
 FT DOMAIN 58 154 PLASTOCYANIN-LIKE.
 FT METAL 94 94 COPPER (BY SIMILARITY).
 FT METAL 139 139 COPPER (BY SIMILARITY).
 FT METAL 142 142 COPPER (BY SIMILARITY).
 FT METAL 147 147 COPPER (BY SIMILARITY).
 SQ SEQUENCE 154 AA; 15577 MW; E45725D5B5F400D CRC64;
 Query Match 63.0%; Score 34; DB 1; Length 154;
 Best Local Similarity 54.5%; Pred. No. 6.8;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EEVVPXGMDYS 11
 Db 100 EDVPSGVDS 110
 RESULT 9
 ID PLAS HORVU STANDARD; PRT; 155 AA.
 AC P08248;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Plastocyanin, chloroplast precursor.
 GN PETE.
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Hordeum.
 OX NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Bomli;
 RA Nielsen O.S., Gausing K.;

RT "The precursor of barley plastocyanin: sequence of cDNA clones and
 RL gene expression in different tissues.";
 RL FEBS Lett. 225:159-162(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. NK 1558;
 RX MEDLINE=94039081; PubMed=8223592;
 RA Nielsen P., Gausing K.;
 RT "In vitro binding of nuclear proteins to the barley plastocyanin gene
 RT promoter region.";
 RL Eur. J. Biochem. 217:97-104(1993).
 CC -!- FUNCTION: Participates in electron transfer between P700 and the
 CC cytochrome b6-f complex in photosystem I.
 CC -!- SUBCELLULAR LOCATION: LOOSELY BOUND TO THE INNER THYLAKOID
 CC MEMBRANE SURFACE IN CHLOROPLASTS.
 CC -!- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; Y00704; CAA68696.1; -;
 DR EMBL; Z28347; CAA82201.1; -;
 DR PIR; S00206; S00206.
 DR HSP; P00289; 2PCF.
 DR InterPro; IPR000923; BlueCu_1.
 DR InterPro; IPR001235; Copper_blue.
 DR Pfam; PF00127; copper-bind; 1.
 DR PRINTS; PR00156; COPPERBLUE.
 DR ProDom; PD001235; Copper blue; 1.
 DR PROSITE; PS00196; COPPER_BLUE; 1.
 KW Chloroplast; Electron transport; Copper; Thylakoid; Membrane;
 KW Transit peptide.
 FT TRANSIT 1 58 CHLOROPLAST.
 FT CHAIN 59 155 PLASTOCYANIN.
 FT DOMAIN 59 155 PLASTOCYANIN-LIKE.
 FT METAL 95 95 COPPER (BY SIMILARITY).
 FT METAL 140 140 COPPER (BY SIMILARITY).
 FT METAL 143 143 COPPER (BY SIMILARITY).
 FT METAL 148 148 COPPER (BY SIMILARITY).
 FT VARIANT 120 120 T -> N (IN CV. NK 1558).
 SQ SEQUENCE 155 AA; 15709 MW; DAA7EABE5F6F4F91 CRC64;
 Query Match 63.0%; Score 34; DB 1; Length 155;
 Best Local Similarity 54.5%; Pred. No. 6.9;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EEVVPXGMDYS 11
 Db 101 EDVPSGVDS 111
 RESULT 10
 ID PLAT POPNI STANDARD; PRT; 168 AA.
 AC P11970;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Plastocyanin B, chloroplast precursor.
 GN PETE.
 OS Populus nigra (lombardy poplar).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Malpighiales; Salicaceae; Populus.
 OX NCBI_TaxID=3691;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Italica; TISSUE=Leaf;

```

RA Reichert J., Jenzelewski V., Haehnel W.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 70-168.
RX STRAIN=cv. Italica;
RA Dmitrova M.I., Egorov C.A., Donchev A.A., Atanasov B.P.;
RT "Complete amino acid sequence of poplar plastocyanin b.";
RL FEBS Lett. 226:17-22(1987).
CC -1- FUNCTION: Participates in electron transfer between P700 and the
CC cytochrome b6-f complex in photosystem I.
CC -1- SUBCELLULAR LOCATION: LOOSELY BOUND TO THE INNER THYLAKOID
CC MEMBRANE SURFACE IN CHLOROPLASTS.
CC -1- SIMILARITY: THERE ARE 12 DIFFERENCES BETWEEN THE SEQUENCE OF
CC POPLAR PLASTOCYANINS A AND B.
CC -1- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z50186; CAA30565.1; -.
CC PIR; S00210; S00210.
CC HSP; P00299; IPLC.
CC InterPro; IPR000923; BlueCu_1.
CC InterPro; IPR001235; Copper_blue.
CC Pfam; PF00127; copper_bind; 1.
CC PRINTS; P00156; COPPERBLUE.
CC PRODOM; PD001235; Copper_blue; 1.
CC PROSITE; PS00196; COPPER_BLUE; 1.
CC Chloroplast; Electron transport; Copper; Thylakoid; Membrane;
KW Transit peptide; Multigene family.
FT TRANSIT 1 69 CHLOROPLAST.
FT CHAIN 70 168 PLASTOCYANIN B.
FT DOMAIN 70 168 PLASTOCYANIN-LIKE.
FT METAL 106 106 COPPER.
FT METAL 153 153 COPPER.
FT METAL 156 156 COPPER.
FT METAL 161 161 COPPER.
SQ SEQUENCE 168 AA; 16981 MW; F20DA6EA2038AEEA CRC64;

Query Match 63.0%; Score 34; DB 1; Length 168;
Best Local Similarity 54.5%; Pred. No. 7.5;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EVVVPXGMDYS 11
Db 112 EDVPSGVDVS 122

RESULT 11
SYDB BORBU
ID -SYDB_BORBU STANDARD; PRT; 566 AA.
AC P94283;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20) (Phenylalanine--
DE tRNA ligase beta chain) (PHERS).
GN PHER OR B00514.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RA Barbour A.G., Hinnebusch J.;
RT "Phenylalanyl-tRNA synthetase genes (alpha and beta subunits) and
RT thioredoxin reductase gene of Borrelia burgdorferi.";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.

```

```

RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Caejens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Uterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RL burgdorferi.";
RN Nature 390:580-586(1997).
CC -1- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA(Phe) = AMP +
CC diphosphate + L-phenylalanyl-tRNA(Phe). TWO BETA CHAINS (BY
CC -1- SUBUNIT: Tetramer of two alpha and two beta chains (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE PHENYLALANYL-tRNA SYNTHETASE BETA
CC CHAIN FAMILY. SUBFAMILY 2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U82978; AAB41019.1; -.
CC EMBL; AE001153; AAC66870.1; -.
CC TIGR; BB0514; -.
CC InterPro; IPR005147; B5.
CC InterPro; IPR004531; PheT_arch.
CC Pfam; PF03484; B5; 1.
CC TIGRFAMS; TIGR00471; pheT arch; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
SQ SEQUENCE 566 AA; 65173 MW; 9D48C8B5D6D3B74B CRC64;

Query Match 63.0%; Score 34; DB 1; Length 566;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 VPXGMDY 10
Db 169 VPFQMDY 175

RESULT 12
Y939_METJA
ID -Y939_METJA STANDARD; PRT; 276 AA.
AC Q58349;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein M70939.
GN M70939.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,

```

RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cottan M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RT jannaschii";
 RL Science 273:1058-1073(1996).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U67537; AAB98946.1; -
 DR TIGR; M0939; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 276 AA; 33454 MW; 97BD69D32BC8FDF CRC64;
 Query Match 61.1%; Score 33; DB 1; Length 276;
 Best Local Similarity 45.5%; Pred. No. 21;
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EHVVPXGMDYS 11
 DB 141 EEIENGMEHS 151
 RESULT 13
 PANC_PSEAE STANDARD; PRT; 283 AA.
 ID PANC_PSEAE
 AC Q9HV69;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Pantoate-beta-alanine ligase (EC 6.3.2.1) (Pantothenate synthetase)
 DE (Pantoate activating enzyme).
 GN PANC OR PA4730.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OC NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PAO1;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Golltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
 RT opportunistic pathogen";
 RL Nature 406:959-964(2000).
 CC -!- CATALYTIC ACTIVITY: ATP + (R)-pantoate + beta-alanine = AMP +
 CC diphosphate + (R)-pantothenate.
 CC -!- PATHWAY: Pantothenate biosynthesis; last step.
 CC -!- SIMILARITY: BELONGS TO THE PANTOTHENATE SYNTHETASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE004886; AAG08116.1; -
 DR InterPro; IPR003721; Pantoate_ligase.
 DR Pfam; PF02569; Pantoate_ligase; 1.
 DR TIGRFAMs; TIGR00018; panC; 1.

KW Pantothenate biosynthesis; Ligase; Complete proteome.
 SQ SEQUENCE 283 AA; 30836 MW; C494949AB40E14E7 CRC64;
 Query Match 61.1%; Score 33; DB 1; Length 283;
 Best Local Similarity 66.7%; Pred. No. 21;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EHVVPXGMD 9
 DB 96 EEMYPDGM 104
 RESULT 14
 HMPA_VIBCH STANDARD; PRT; 394 AA.
 ID HMPA_VIBCH
 AC Q9KMY3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 13-JUN-2002 (Rel. 41, Last annotation update)
 DE Flavohemoprotein (Hemoglobin-like protein) (Flavohemoglobin).
 GN HMP OR VCA0183.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OC NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=El Tor N16961 / Serotype O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 RT cholerae";
 RL Nature 406:477-483(2000).
 CC -!- DOMAIN: CONSISTS OF TWO DISTINCT DOMAINS; ONE IS A HEME-CONTAINING
 CC OXYGEN BINDING DOMAIN IN THE N-TERMINAL REGION AND THE OTHER IS AN
 CC FAD-CONTAINING REDUCTASE DOMAIN FOUND IN THE C-TERMINAL REGION.
 CC -!- SIMILARITY: BELONGS TO THE GLOBIN FAMILY. TWO-DOMAIN
 CC FLAVOHEMOPROTEINS SUBFAMILY.
 CC -!- SIMILARITY: TO A NUMBER OF FAD/NAD(P) FLAVOPROTEIN
 CC OXIDOREDUCTASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE004358; AAF96096.1; -
 DR HSSP; P39662; 1CQX.
 DR TIGR; VCA0183; -
 DR InterPro; IPR001834; Cyt_B5_reductase.
 DR InterPro; IPR001709; FPN_cyt_reductase.
 DR InterPro; IPR000971; Globin.
 DR InterPro; IPR001433; Oxid_FAD/NAD(P).
 DR InterPro; IPR001221; Phe_hydroxylase.
 DR Pfam; PF00042; globin; 1.
 DR Pfam; PF00175; NAD_binding; 1.
 DR Pfam; PF00970; FAD_binding_6; 1.
 DR PRINTS; PR00371; PFNCR.
 DR PROSITE; PR00410; PHEHYDRLASE.
 DR PROSITE; PS01033; GLOBIN; 1.
 DR Oxidoreductase; NADP; Heme; Flavoprotein; FAD; Iron transport;
 KW Oxygen transport; Transport; Complete proteome.
 FT DOMAIN 1 136 GLOBIN.
 FT METAL 53 53 IRON (HEME DISTAL LIGAND)
 FT (BY SIMILARITY).

FT METAL 85 85 IRON (HEME PROXIMAL LIGAND)
 FT NP BIND 268 273 NADP (RIBOSE PART) (BY SIMILARITY).
 SQ SEQUENCE 394 AA; 44191 MW; DDA3490FAE28823A CRC64;

Query Match 61.1%; Score 33; DB 1; Length 394;
 Best Local Similarity 66.7%; Pred. No. 30;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 EVVVPXGMDY 10
 |||||
 Db 194 EVTPEGSY 202

RESULT 15

ACDM RAT STANDARD; PRT; 421 AA.
 AC P08503;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Acyl-CoA dehydrogenase, medium-chain specific, mitochondrial precursor
 DE (EC 1.3.99.3) (MCAD).
 GN ACADM.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Liver;
 RX MEDLINE=87280028; PubMed=3611054;
 RA Matsubara Y., Kraus J.P., Ozasa H., Glassberg R., Finocchiaro G.,
 RA Ikeda Y., Mole J., Rosenberg L.E., Tanaka K.
 RT "Molecular cloning and nucleotide sequence of cDNA encoding the
 RT entire precursor of rat liver medium chain acyl coenzyme A
 RT dehydrogenase."
 RL J. Biol. Chem. 262:10104-10108(1987).
 CC -!- FUNCTION: THIS ENZYME IS SPECIFIC FOR ACYL CHAIN LENGTHS OF 4 TO
 CC 16.
 CC -!- CATALYTIC ACTIVITY: Acyl-CoA + ETF = 2,3-dehydroacyl-CoA + reduced
 CC ETF.
 CC -!- COFACTOR: FAD.
 CC -!- PATHWAY: Mitochondrial fatty acid beta-oxidation system; first
 CC step.
 CC -!- SUBUNIT: HOMOTETRAMER.
 CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -!- MISCELLANEOUS: A NUMBER OF STRAIGHT-CHAIN ACYL-COA DEHYDROGENASES
 CC OF DIFFERENT SUBSTRATE SPECIFICITIES ARE PRESENT IN MAMMALIAN
 CC TISSUES.
 CC -!- SIMILARITY: BELONGS TO THE ACYL-COA DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; J02791; AAA0670.1; -.
 CC PIR; A28436; DERTCN.
 CC HSP; P11310; LEGD.
 CC InterPro; IPR001552; Acyl-CoA dh.
 CC Pfam; PF00441; Acyl-CoA dh; 1.
 CC Pfam; PF02770; Acyl-CoA dh_M; 1.
 CC Pfam; PF02771; Acyl-CoA dh_N; 1.
 CC PROSITE; PS00072; ACYL COA DH 1; 1.
 CC PROSITE; PS00073; ACYL COA DH 2; 1.
 CC Oxidoreductase; Flavoprotein; FAD; Fatty acid metabolism;
 CC Mitochondrion; Transit peptide.
 CC TRANSIT 1 25 MITOCHONDRION
 CC CHAIN 26 421 ACYL-COA DEHYDROGENASE, MEDIUM-CHAIN

FT ACT_SITE 193 193 SPECIFIC.
 FT N(5) OF THE FAD COFACTOR (BY SIMILARITY).
 FT BASE (BY SIMILARITY).
 SQ SEQUENCE 421 AA; 46555 MW; 2CF076F8C919BDE8 CRC64;
 Query Match 61.1%; Score 33; DB 1; Length 421;
 Best Local Similarity 50.0%; Pred. No. 32;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 EEVVPXGMDY 10
 |||||
 Db 58 EEIIPVADY 67

Search completed: June 4, 2003, 13:11:46
 Job time : 7.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:08:49 ; Search time 11 Seconds
(without alignments)
96.134 Million cell updates/sec

Title: AUDET-909-3

Perfect score: 54

Sequence: 1 eevvpXgmndys 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	74.1	156	2 S54619	hypothetical prote
2	38	70.4	363	2 D69551	conserved hypothet
3	36	66.7	102	2 A42452	V1 protein - tobac
4	36	66.7	341	2 S72649	sucrose-phosphate
5	36	66.7	348	2 S72650	sucrose-phosphate
6	36	66.7	460	2 G96764	unknown protein F2
7	36	66.7	1049	2 JCA783	sucrose-phosphate
8	36	66.7	1068	1 JQ1329	sucrose-phosphate
9	36	66.7	1081	2 TQ9837	sucrose-phosphate
10	36	66.7	1083	2 TQ4062	sucrose-phosphate
11	36	66.7	1084	2 TQ4103	sucrose-phosphate
12	35	64.8	425	2 T24111	hypothetical prote
13	35	64.8	433	2 H87660	peptidoglycan-bind
14	35	64.8	440	2 H72784	probable alkaline
15	35	64.8	1150	2 T20173	hypothetical prote
16	35	64.8	1474	2 F69009	probable membrane
17	35	64.8	2747	2 B49132	fat facets (faf) s
18	34	63.0	99	2 S00210	plastocyanin b - L
19	34	63.0	155	2 S38255	plastocyanin b pre
20	34	63.0	168	2 S58208	plastocyanin b pre
21	34	63.0	290	2 D98182	O6-methylguanidine-D
22	34	63.0	290	2 AG3104	6-O-methylguanidine
23	34	63.0	296	2 F72745	hypothetical prote
24	34	63.0	357	1 G69290	probable hexosyltr
25	34	63.0	366	2 G69350	L-lactate dehydrog
26	34	63.0	565	2 E86665	ABC transporter AT
27	34	63.0	566	2 A70164	phenylalanine-tRNA
28	34	63.0	587	2 F81138	succinate dehydrog
29	34	63.0	906	2 T48898	disease resistance

RESULT 1

S54619

hypothetical protein YOR013w - yeast (Saccharomyces cerevisiae)

N/Alternate names: hypothetical protein O2612; hypothetical protein YOL303.3

C/Species: Saccharomyces cerevisiae

C/Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 19-Apr-2002

C/Accession: S54619; S66879

R/de Haan, M.; Maarse, A.C.; Grivell, L.A.

submitted to the EMBL Data Library, May 1995

A/Reference number: S54617

A/Accession: S54619

A/Molecule type: DNA

A/Residues: 1-156 <DEH>

A/Cross-references: EMBL:X87331; NID:g1041652; PIDN:CAA60762.1; PID:g829123

R/de Haan, M.; Grivell, L.A.; Maarse, A.C.

submitted to the Protein Sequence Database, July 1996

A/Reference number: S66877

A/Accession: S66879

A/Molecule type: DNA

A/Residues: 1-156 <DEW>

A/Cross-references: EMBL:Z74920; NID:g1420109; PIDN:CAA99201.1; PID:g1420111; MIPS:YOR013

A/Experimental source: strain S288C

C/Genetics:

A/Cross-references: SGD:S0005539

A/Map position: 15R

C/Superfamily: hypothetical protein YOR013w

Query Match 74.1%; Score 40; DB 2; Length 156;

Best Local Similarity 77.8%; Pred. No. 1;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EYVPXGMDY 10

Db 50 EYVPLGMDY 58

||:|||||

||:|||||

RESULT 2

D69551

conserved hypothetical protein AF2411 - Archaeoglobus fulgidus

C/Species: Archaeoglobus fulgidus

C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999

C/Accession: D69551

R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson,

.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A/Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.H.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo

A/Reference number: A69250; MUID:98049343; PMID:9389475

A/Accession: D69551

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-363 <KLE>
A:Cross-references: GB:AE001109; GB:AE000782; NID:g2689432; PIDN:AAB91255.1; PID:g265068

Query Match 70.4%; Score 38; DB 2; Length 363;
Best Local Similarity 54.5%; Pred. No. 6.8;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVVVPXGMDYS 11
|:|:|:|:|:
Db 120 ENIVPGIDFS 130

RESULT 3
A42452
V1 protein - tobacco yellow dwarf virus (strain Australia)
C:Species: tobacco yellow dwarf virus
C:Date: 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change 08-Oct-1999
C:Accession: A42452
R:Morris, B.A.M.; Richardson, K.A.; Haley, A.; Zhan, X.; Thomas, J.E.
Virology 187, 633-642, 1992
A:Title: The nucleotide sequence of the infectious cloned DNA component of tobacco yellow
A:Reference number: A42452; MUID:92188538; PMID:1546458
A:Accession: A42452
A:Molecule type: DNA
A:Residues: 1-102 <MOR>
A:Cross-references: GB:M81103; NID:g335283; PIDN:AAA47947.1; PID:g335284

Query Match 66.7%; Score 36; DB 2; Length 102;
Best Local Similarity 60.0%; Pred. No. 4.2;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EVVVPXGMDYS 11
|:|:|:|:|:
Db 7 QVVPISGINS 16

RESULT 4
S72649
sucrose-phosphate synthase (EC 2.4.1.14) isoform 2 - Citrus unshiu (fragment)
C:Species: Citrus unshiu
C:Date: 24-Oct-1998 #sequence_revision 24-Oct-1998 #text_change 21-Jul-2000
C:Accession: S72649
R:Komatsu, A.; Takanokura, Y.; Omura, M.; Akihama, T.
Mol. Gen. Genet. 252, 346-351, 1996
A:Title: Cloning and molecular analysis of cDNAs encoding three sucrose phosphate synthase
A:Reference number: S72648; MUID:96439842; PMID:8842155
A:Accession: S72649
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-341 <KOM>
A:Cross-references: EMBL:AB006319; NID:g2588891; PIDN:BAA23215.1; PID:g2588892
A:Experimental source: fruit, cv. Miyagawa-Wase
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1997
C:Genetics:
A:Gene: SPS2
C:Function:
A:Description: catalyzes formation of sucrose-6-phosphate from UDPglucose and D-fructose
A:Pathway: sucrose biosynthesis
A:Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homology
C:Keywords: glycosyltransferase; hexosyltransferase; sucrose biosynthesis
F:1-341/Domain: sucrose/sucrose-phosphate synthase homology (fragment) <SSPS>

Query Match 66.7%; Score 36; DB 2; Length 341;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VVPXGMDYS 11
|:|:|:|:|:
Db 228 VPPGMDFS 236

RESULT 5
S72650
sucrose-phosphate synthase (EC 2.4.1.14) isoform 3 - Citrus unshiu (fragment)
C:Species: Citrus unshiu
C:Date: 24-Oct-1998 #sequence_revision 24-Oct-1998 #text_change 21-Jul-2000
C:Accession: S72650
R:Komatsu, A.; Takanokura, Y.; Omura, M.; Akihama, T.
Mol. Gen. Genet. 252, 346-351, 1996
A:Title: Cloning and molecular analysis of cDNAs encoding three sucrose phosphate synthase
A:Reference number: S72648; MUID:96439842; PMID:8842155
A:Accession: S72650
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-348 <KOM>
A:Cross-references: EMBL:AB006660; NID:g2351059; PIDN:BAA22071.1; PID:g2351060
A:Experimental source: fruit, cv. Miyagawa-Wase
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1997
C:Genetics:
A:Gene: SPS3
C:Function:
A:Description: catalyzes formation of sucrose-6-phosphate from UDPglucose and D-fructose
A:Pathway: sucrose biosynthesis
A:Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homology
C:Keywords: glycosyltransferase; hexosyltransferase; sucrose biosynthesis
F:1-348/Domain: sucrose/sucrose-phosphate synthase homology (fragment) <SSPS>

Query Match 66.7%; Score 36; DB 2; Length 348;
Best Local Similarity 66.7%; Pred. No. 17;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VVPXGMDYS 11
|:|:|:|:|:
Db 234 VPPGMDFS 242

RESULT 6
G96764
unknown protein F25P22.17 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: G96764
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
anssen, N.F.; Hughes, B.; Huizlar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Saizberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: G96764
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-460 <STO>
A:Cross-references: GB:AE005173; NID:g6692750; PIDN:AAF24856.1; GSPD:GN00141
C:Genetics:
A:Gene: F25P22.17
A:Map position: 1

Query Match 66.7%; Score 36; DB 2; Length 460;
Best Local Similarity 70.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 BEVVPXGMDY 10
|:|:|:|:|:
Db 218 EEDVPSANDY 227

RESULT 7
JC4783
sucrose-phosphate synthase (EC 2.4.1.14) - rice
C:Species: Oryza sativa (rice)

C>Date: 10-May-1996 #sequence_revision 16-Aug-1996 #text_change 18-Jun-1999
 C/Accession: J04783
 R;Valdez-Alarcon, J.J.; Ferrando, M.; Salerno, G.; Jimenez-Moraila, B.; Herrera-Estrella
 Gene 170, 217-222, 1996
 A>Title: Characterization of a rice sucrose-phosphate synthase-encoding gene.
 A/Reference number: J04783; MUID:96235138; PMID:8666248
 A/Accession: J04783
 A/Molecule type: mRNA
 A/Residues: 1-1049 <VAL>
 A/Cross-references: GB:U33175; NID:g1449931; PIDN:AA049379.1; PID:g988270
 A/Note: UDPglucose-fructose-phosphate glucosyltransferase; Sucrosephosphate-UDPglucosyltr
 C/Comment: This enzyme catalyzes the formation of sucrose-phosphate form UDP-glucose and
 C/Genetics:
 A/Gene: Sp81
 A/Introns: 24/1; 103/3; 183/3; 205/3; 435/3; 475/3; 519/3; 578/3; 596/3; 617/3; 931/3; 9
 C/Function:
 A/Description: catalyzes the formation of sucrose-6-phosphate from UDPglucose and D-fruc
 A/Pathway: sucrose biosynthesis
 C/Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homology
 C/Keywords: glycosyltransferase; hexosyltransferase; sucrose biosynthesis
 F;180-663/Domain: sucrose/sucrose-phosphate synthase homology <SSPS>

Query Match 66.7%; Score 36; DB 2; Length 1049;
 Best Local Similarity 66.7%; Pred. No. 58;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
 Db 436 VIPPGMDFS 444
 |:|:|:|:|

RESULT 8
 J01329
 sucrose-phosphate synthase (EC 2.4.1.14) - maize
 C/Species: Zea mays (maize)
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C/Accession: J01329; PQ0260
 R;Worrell, A.C.; Bruneau, J.M.; Summerfelt, K.; Boersig, M.; Voelker, T.A.
 Plant Cell 3, 1121-1130, 1991
 A>Title: Expression of a maize sucrose phosphate synthase in tomato alters leaf carbonyd
 A/Reference number: J01329; MUID:92338837; PMID:1840396
 A/Accession: J01329
 A/Molecule type: mRNA
 A/Residues: 1-1068 <WOR>
 A/Cross-references: GB:M97550; NID:g168625; PIDN:AAA33513.1; PID:g168626
 A/Accession: PQ0260
 A/Molecule type: protein
 A/Residues: 71-74;206-212;471-481;872-892 <WOR1>
 C/Comment: This enzyme transfers the glucosyl group from UDPglucose to fructose-6-phosph
 C/Comment: This enzyme is involved in the regulation of carbon partitioning in the leave
 C/Function:
 A/Description: catalyzes the formation of sucrose-6-phosphate from UDPglucose and D-fruc
 A/Pathway: sucrose biosynthesis
 C/Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homology
 C/Keywords: glycosyltransferase; hexosyltransferase; sucrose biosynthesis
 F;178-666/Domain: sucrose/sucrose-phosphate synthase homology <SSPS>

Query Match 66.7%; Score 36; DB 1; Length 1068;
 Best Local Similarity 66.7%; Pred. No. 59;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
 Db 435 VIPPGMDFS 443
 |:|:|:|:|

RESULT 9
 T09837
 sucrose-phosphate synthase (EC 2.4.1.14) isoform 2 - Craterostigma plantagineum
 C/Species: Craterostigma plantagineum
 C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
 C/Accession: T09837
 R;Ingram, J.; Chandler, J.W.; Gallegher, L.; Salamini, F.; Bartels, D.

Plant Physiol. 115, 113-121, 1997
 A>Title: Analysis of cDNA clones encoding sucrose-phosphate synthase in relation to suga
 A/Reference number: Z16874; MUID:97451773; PMID:9306694
 A/Accession: T09837
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-1081 <ING>
 A/Cross-references: EMBL:Y11795; NID:g2190349; PIDN:CAA72491.1; PID:g2190350
 A/Experimental source: ABA-treated callus
 C/Genetics:
 A/Gene: sps2
 C/Function:
 A/Description: catalyzes the formation of sucrose-6-phosphate from UDPglucose and D-fruct
 A/Pathway: sucrose biosynthesis
 C/Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homology
 C/Keywords: glycosyltransferase; hexosyltransferase; sucrose biosynthesis
 F;176-674/Domain: sucrose/sucrose-phosphate synthase homology <SSS>

Query Match 66.7%; Score 36; DB 2; Length 1081;
 Best Local Similarity 66.7%; Pred. No. 60;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
 Db 445 VIPPGMDFS 453
 |:|:|:|:|

RESULT 10
 T04062
 sucrose-phosphate synthase homolog F28M11.40 - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 16-Jul-1999
 C/Accession: T04062
 R;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.
 submitted to the Protein Sequence Database, March 1999
 A/Reference number: Z15184
 A/Accession: T04062
 A/Molecule type: DNA
 A/Residues: 1-1083 <BEV>
 A/Cross-references: EMBL:AL049487
 A/Experimental source: cultivar Columbia; BAC clone F28M11
 C/Genetics:
 A/Map position: 4
 A/Introns: 86/3; 116/3; 255/3; 322/2; 482/3; 526/3; 570/3; 629/3; 647/3; 668/3; 949/3; 9;
 A/Note: F28M11.40
 C/Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homology
 F;230-714/Domain: sucrose/sucrose-phosphate synthase homology <SSPS>

Query Match 66.7%; Score 36; DB 2; Length 1083;
 Best Local Similarity 66.7%; Pred. No. 60;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
 Db 483 VIPPGMDFS 491
 |:|:|:|:|

RESULT 11
 T04103
 sucrose-phosphate synthase (EC 2.4.1.14) 1 - rice
 C/Species: Oryza sativa (rice)
 C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 20-Jun-2000
 C/Accession: T04103
 R;Sakamoto, M.; Satozawa, T.; Kishimoto, N.; Higo, K.; Shinada, H.; Fujimura, T.
 Plant Sci. 112, 207-217, 1995
 A>Title: Structure and RFLP mapping of a rice sucrose phosphate synthase (SPS) gene that
 A/Reference number: Z15212
 A/Accession: T04103
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1084 <SAK>
 A/Cross-references: EMBL:D45890; PIDN:BAA08304.1
 A/Experimental source: subsp. Japonica

C:Genetics:
A:Gene: Spal
A:Map position: 1
A:Introns: 120/3; 200/2; 221/3; 452/3; 492/3; 536/3; 595/3; 613/3; 634/3; 946/3; 989/2
C:Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homology
C:Keywords: Glycosyltransferase; hexosyltransferase
F:196-680/Domain: sucrose-phosphate synthase homology <SSPS>

Query Match 66.7%; Score 36; DB 2; Length 1084;
Best Local Similarity 66.7%; Pred. No. 60;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VVFXGMDYS 11
|:|:|:|:
Db 453 VVPPGMDFS 461

RESULT 12
T24111
hypothetical protein R10D12.10 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T24111
R:Percy, C.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19842
A:Accession: T24111
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-425 <WIL>
A:Cross-references: EMBL:Z81109; PIDN:CAB03241.1; GSPDB:GN00023; CESP:R10D12.10
A:Experimental source: clone R10D12
C:Genetics:
A:Gene: CESP:R10D12.10
A:Map position: 5
A:Introns: 23/3; 56/3; 113/3; 257/2

Query Match 64.8%; Score 35; DB 2; Length 425;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EEVVPXGMDY 10
|:|:|:|:
Db 335 EQIVPGGLQY 344

RESULT 13
H87660
peptidoglycan-binding protein, probable [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: H87660
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: H87660
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-433 <STO>
A:Cross-references: GB:AB005673; NID:gl3425020; PIDN:AAK25284.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC322

Query Match 64.8%; Score 35; DB 2; Length 433;
Best Local Similarity 54.5%; Pred. No. 34;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMDYS 11
|:|:|:|:
Db 266 EVILPFGDYS 276

RESULT 14

H72784
probable alkaline proteinase APE0263 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: H72784

R:Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takaha
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ki
DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: H72784

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-440 <KAW>

A:Cross-references: DDBJ:AP000058; NID:G5103388; PIDN:BAA79178.1; PID:G5103657

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE0263

C:Superfamily: subtilisin; subtilisin homology

Query Match 64.8%; Score 35; DB 2; Length 440;
Best Local Similarity 66.7%; Pred. No. 35;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EVVFXGMDY 10
|:|:|:|:
Db 120 EVLPWGVY 128

RESULT 15

T20173

hypothetical protein CS3A5.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999

C:Accession: T20173; T23857

R:Mortimore, B.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19232

A:Accession: T20173

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-1150 <WIL>

A:Cross-references: EMBL:Z81486; PIDN:CAB03994.1; GSPDB:GN00023; CESP:CS3A5.2

A:Experimental source: clone CS3A5

R:Matthews, L.

submitted to the EMBL Data Library, August 1996

A:Reference number: Z19808

A:Accession: T23857

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-1150 <WIL>

A:Cross-references: EMBL:Z78015; PIDN:CAB01437.1; GSPDB:GN00023; CESP:CS3A5.2

A:Experimental source: clone R02D5

C:Genetics:

A:Gene: CESP:CS3A5.2

A:Map position: 5

A:Introns: 33/3; 63/3; 132/3; 169/3; 221/3; 299/3; 379/2; 423/2; 438/2; 471/1; 513/2; 656

Query Match 64.8%; Score 35; DB 2; Length 1150;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VVFXGMDYS 11
|:|:|:|:
Db 562 VLPVGDYS 570

Search completed: June 4, 2003, 13:14:16
Job time : 13 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:13:35 ; Search time 14.25 Seconds
(without alignments)
79.694 Million cell updates/sec

Title: AUDET-909-3

Perfect score: 54

Sequence: 1 eevpxgmdys 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 101240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/PCT NEW PUB. pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT NEW PUB. pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW PUB. pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB. pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW PUB. pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB. pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB. pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB. pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW PUB. pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB. pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW PUB. pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB. pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW PUB. pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB. pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	66.7	1049	9	US-10-217-700-10
2	36	66.7	1068	9	US-10-217-700-8
3	36	66.7	1081	9	US-10-217-700-4
4	36	66.7	1083	9	US-10-217-700-11
5	36	66.7	1084	9	US-10-217-700-9
6	35	64.8	440	9	US-09-813-408-27
7	34	63.0	1062	10	US-09-815-242-5111
8	34	63.0	3472	9	US-10-027-806-4
9	34	63.0	3472	9	US-10-034-623-4
10	34	63.0	3472	9	US-10-027-801-4
11	33	61.1	59	10	US-09-948-080-14
12	33	61.1	283	9	US-09-738-626-4881
13	33	61.1	239	10	US-09-815-242-10697
14	33	61.1	736	9	US-09-978-295A-526
15	33	61.1	736	9	US-09-978-697-526
16	33	61.1	736	9	US-09-978-192A-526
17	33	61.1	736	9	US-09-999-832A-526
18	33	61.1	736	9	US-09-978-189-526
19	33	61.1	736	9	US-10-174-590-420

ALIGNMENTS

RESULT 1

US-10-217-700-10
; Sequence 10, Application US/10217700
; Publication No. US20030070191A1
; GENERAL INFORMATION:
; APPLICANT: Haigler, Candace H.
; APPLICANT: Holaday, A. Scott
; TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED
; FILE OF INVENTION: EXPRESSION OF SUCROSE PHOSPHATE SYNTHASE
; FILE REFERENCE: 201304/1000
; CURRENT APPLICATION NUMBER: US/10/217,700
; CURRENT FILING DATE: 2002-08-12
; EARLIER APPLICATION NUMBER: 09/394,272
; EARLIER FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-217-700-10

Query Match 66.7%; Score 36; DB 9; Length 1049;
Best Local Similarity 66.7%; Pred. No. 96;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VVEXGMDYS 11

Db 436 VVPPGMDFS 444

RESULT 2

US-10-217-700-8
; Sequence 8, Application US/10217700
; Publication No. US20030070191A1
; GENERAL INFORMATION:
; APPLICANT: Haigler, Candace H.
; APPLICANT: Holaday, A. Scott
; TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED
; FILE OF INVENTION: EXPRESSION OF SUCROSE PHOSPHATE SYNTHASE
; FILE REFERENCE: 201304/1000
; CURRENT APPLICATION NUMBER: US/10/217,700

; CURRENT FILING DATE: 2002-08-12
; EARLIER APPLICATION NUMBER: 09/394,272
; EARLIER FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1068
; TYPE: PRT
; ORGANISM: Zea mays
US-10-217-700-8

Query Match 66.7%; Score 36; DB 9; Length 1068;
Best Local Similarity 66.7%; Pred. No. 98;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
|:|:|:|:|:
Db 435 VIPPGMDFS 443

RESULT 3

US-10-217-700-4
; Sequence 4, Application US/10217700
; Publication No. US20030070191A1
; GENERAL INFORMATION:
; APPLICANT: Haigler, Candace H.
; APPLICANT: Holaday, A. Scott
; TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED
; FILE REFERENCE: 201304/1000
; CURRENT APPLICATION NUMBER: US/10/217,700
; CURRENT FILING DATE: 2002-08-12
; EARLIER APPLICATION NUMBER: 09/394,272
; EARLIER FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1081
; TYPE: PRT
; ORGANISM: Cratogeomys plantagineum
US-10-217-700-4

Query Match 66.7%; Score 36; DB 9; Length 1081;
Best Local Similarity 66.7%; Pred. No. 1e-02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
|:|:|:|:|:
Db 445 VIPPGMDFS 453

RESULT 4

US-10-217-700-11
; Sequence 11, Application US/10217700
; Publication No. US20030070191A1
; GENERAL INFORMATION:
; APPLICANT: Haigler, Candace H.
; APPLICANT: Holaday, A. Scott
; TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED
; FILE REFERENCE: 201304/1000
; CURRENT APPLICATION NUMBER: US/10/217,700
; CURRENT FILING DATE: 2002-08-12
; EARLIER APPLICATION NUMBER: 09/394,272
; EARLIER FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 1083
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-217-700-11

Query Match 66.7%; Score 36; DB 9; Length 1083;
Best Local Similarity 66.7%; Pred. No. 1e-02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
|:|:|:|:|:
Db 483 VIPPGMDFS 491

RESULT 5

US-10-217-700-9
; Sequence 9, Application US/10217700
; Publication No. US20030070191A1
; GENERAL INFORMATION:
; APPLICANT: Haigler, Candace H.
; APPLICANT: Holaday, A. Scott
; TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED
; FILE REFERENCE: 201304/1000
; CURRENT APPLICATION NUMBER: US/10/217,700
; CURRENT FILING DATE: 2002-08-12
; EARLIER APPLICATION NUMBER: 09/394,272
; EARLIER FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1084
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-217-700-9

Query Match 66.7%; Score 36; DB 9; Length 1084;
Best Local Similarity 66.7%; Pred. No. 1e-02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
|:|:|:|:|:
Db 453 VIPPGMDFS 461

RESULT 6

US-09-813-408-27
; Sequence 27, Application US/09813408
; Publication No. US20030049619A1
; GENERAL INFORMATION:
; APPLICANT: Delagrave, Simon
; APPLICANT: Marrs, Barry
; TITLE OF INVENTION: Methods For The Synthesis Of Polynucleotides And Combinatorial Li
; FILE REFERENCE: HER0041
; CURRENT APPLICATION NUMBER: US/09/813,408
; CURRENT FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Aeropyrum pernix
US-09-813-408-27

Query Match 64.8%; Score 35; DB 9; Length 440;
Best Local Similarity 66.7%; Pred. No. 58;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMDY 10
|:|:|:|:|:
Db 120 EVLPWGVGY 128

RESULT 7

US-09-815-242-5111
; Sequence 5111, Application US/09815242
; Patent No. US20020061569A1

; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Karl L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; TITLE OF INVENTION: Prokaryotes
 ; FILE REFERENCE: ELITRA.011a
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 5111
 ; LENGTH: 1062
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-815-242-5111

Query Match 63.0%; Score 34; DB 10; Length 1062;
 Best Local Similarity 85.7%; Pred. No. 2.4e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PXGMDYS 11
 | | | | |
 Db 321 PQGMDYS 327

RESULT 8
 US-10-027-806-4
 ; Sequence 4, Application US/10027806
 ; Patent No. US20020160476A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Swanson, Ronald V.
 ; APPLICANT: Feldman, Robert A.
 ; APPLICANT: Schleper, Christa
 ; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
 ; FILE REFERENCE: DCCP.002A
 ; CURRENT APPLICATION NUMBER: US/10/027,806
 ; CURRENT FILING DATE: 2001-12-21
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
 ; NUMBER OF SEQ ID NOS: 123
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 3472
 ; TYPE: PRT
 ; ORGANISM: Cenarchaeum symbiosum
 US-10-027-806-4

Query Match 63.0%; Score 34; DB 9; Length 3472;
 Best Local Similarity 45.5%; Pred. No. 9.1e+02;
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
 | | | | |

Db 2294 EDVIPRGISFS 2304
 RESULT 9
 US-10-034-623-4
 ; Sequence 4, Application US/10034623
 ; Publication No. US20020198365A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Swanson, Ronald V.
 ; APPLICANT: Feldman, Robert A.
 ; APPLICANT: Schleper, Christa
 ; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
 ; FILE REFERENCE: DCCP.002A
 ; CURRENT APPLICATION NUMBER: US/10/034,623
 ; CURRENT FILING DATE: 2001-12-21
 ; PRIOR APPLICATION NUMBER: 09/408,020
 ; PRIOR FILING DATE: 1999-09-29
 ; PRIOR APPLICATION NUMBER: 60/102,294
 ; PRIOR FILING DATE: 1998-09-29
 ; NUMBER OF SEQ ID NOS: 123
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 3472
 ; TYPE: PRT
 ; ORGANISM: Cenarchaeum symbiosum
 US-10-034-623-4

Query Match 63.0%; Score 34; DB 9; Length 3472;
 Best Local Similarity 45.5%; Pred. No. 9.1e+02;
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
 | | | | |
 Db 2294 EDVIPRGISFS 2304

RESULT 10
 US-10-027-801-4
 ; Sequence 4, Application US/10027801
 ; Publication No. US20030054364A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Swanson, Ronald V.
 ; APPLICANT: Feldman, Robert A.
 ; APPLICANT: Schleper, Christa
 ; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
 ; FILE REFERENCE: DCCP.002A
 ; CURRENT APPLICATION NUMBER: US/10/027,801
 ; CURRENT FILING DATE: 2001-12-21
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
 ; NUMBER OF SEQ ID NOS: 123
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 3472
 ; TYPE: PRT
 ; ORGANISM: Cenarchaeum symbiosum
 US-10-027-801-4

Query Match 63.0%; Score 34; DB 9; Length 3472;
 Best Local Similarity 45.5%; Pred. No. 9.1e+02;
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
 | | | | |
 Db 2294 EDVIPRGISFS 2304

RESULT 11
 US-09-948-080-14
 ; Sequence 14, Application US/09948080
 ; Patent No. US20020102702A1
 ; GENERAL INFORMATION:
 ; APPLICANT: VAN DER OSTEN, CLAUS

APPLICANT: HALKIER, TORDEN
APPLICANT: ANDERSEN, CARSTEN
APPLICANT: BAUDITZ, PETER
APPLICANT: HANSEN, PETER KAMP
TITLE OF INVENTION: PROTEASE VARIANTS AND COMPOSITIONS
FILE REFERENCE: 4946,200-US
CURRENT APPLICATION NUMBER: US/09/948,080
PRIOR FILING DATE: 2001-09-06
PRIOR APPLICATION NUMBER: US/08/363,851
PRIOR FILING DATE: 1997-11-04
NUMBER OF SEQ ID NOS: 35
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 14
LENGTH: 59
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-948-080-14

Query Match 61.1%; Score 33; DB 10; Length 59;
Best Local Similarity 45.5%; Pred. No. 15;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EEVVPXGMDYS 11
Db 38 EXHIPGGLEYS 48
: : : : :
: : : : :
: : : : :

RESULT 12

US-09-738-626-4881
Sequence 4881, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 4881
LENGTH: 283
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-4881

Query Match 61.1%; Score 33; DB 9; Length 283;
Best Local Similarity 75.0%; Pred. No. 88;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 VPXGMDYS 11
Db 56 VPAGADYS 63
: : : : :
: : : : :
: : : : :

RESULT 13

US-09-815-242-10697
Sequence 10697, Application US/09815242
Patent No. US20020061569A1

GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10697
LENGTH: 299
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-815-242-10697

Query Match 61.1%; Score 33; DB 10; Length 299;
Best Local Similarity 40.0%; Pred. No. 94;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EEVVPXGMDY 10
Db 218 EQITPTGIEY 227
: : : : :
: : : : :
: : : : :

RESULT 14

US-09-978-295A-526
Sequence 526, Application US/09978295A
Patent No. US20020156006A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnovers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas P.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.

APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2630PIC11
CURRENT APPLICATION NUMBER: US/09/978,295A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598

; PRIOR FILING DATE: 1998-05-07
 ; PRIOR APPLICATION NUMBER: 60/084600
 ; PRIOR FILING DATE: 1998-05-07
 ; PRIOR APPLICATION NUMBER: 60/084627
 ; PRIOR FILING DATE: 1998-05-07
 ; PRIOR APPLICATION NUMBER: 60/084643
 ; PRIOR FILING DATE: 1998-05-07
 ; PRIOR APPLICATION NUMBER: 60/085339
 ; PRIOR FILING DATE: 1998-05-13
 ; PRIOR APPLICATION NUMBER: 60/085338
 ; PRIOR FILING DATE: 1998-05-13
 ; PRIOR APPLICATION NUMBER: 60/085323
 ; PRIOR FILING DATE: 1998-05-13
 ; PRIOR APPLICATION NUMBER: 60/085582
 ; PRIOR FILING DATE: 1998-05-15
 ; PRIOR APPLICATION NUMBER: 60/085700
 ; PRIOR FILING DATE: 1998-05-15
 ; PRIOR APPLICATION NUMBER: 60/085689
 ; PRIOR FILING DATE: 1998-05-15
 ; PRIOR APPLICATION NUMBER: 60/085579
 ; PRIOR FILING DATE: 1998-05-15
 ; PRIOR APPLICATION NUMBER: 60/085580
 ; PRIOR FILING DATE: 1998-05-15
 ; PRIOR APPLICATION NUMBER: 60/085573
 ; PRIOR FILING DATE: 1998-05-15
 ; PRIOR APPLICATION NUMBER: 60/085704
 ; PRIOR FILING DATE: 1998-05-15
 ; PRIOR APPLICATION NUMBER: 60/085697

Query Match 61.1%; Score 33; DB 9; Length 736;

Best Local Similarity 70.0%; Pred. No. 2.6e+02; Mismatches 0; Gaps 0; Indels 3; Indels 0; Gaps 0;

Qy 1 BEVPXGMDY 10
 Db 331 EPVVVXGMDY 340

RESULT 15

; Sequence 526, Application US/09978697
 ; Patent No. US20020169284A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ashkenazi, Avi
 ; APPLICANT: Baker Kevin P.
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan
 ; APPLICANT: Ferrara, Napoleon
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Hillan, Kenneth J.
 ; APPLICANT: Kljavin, Ivar J.
 ; APPLICANT: Kuo, Sophia S.
 ; APPLICANT: Napier, Mary A.
 ; APPLICANT: Pan, James;
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Shelton, David L.
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; TITLE OF INVENTION: Acids Encoding the Same
 ; FILE REFERENCE: P2630P1C27

; CURRENT APPLICATION NUMBER: US/09/978,697
 ; CURRENT FILING DATE: 2001-10-16
 ; PRIOR APPLICATION NUMBER: 09/918585
 ; PRIOR FILING DATE: 2001-07-30
 ; PRIOR APPLICATION NUMBER: 60/062250
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/064249
 ; PRIOR FILING DATE: 1997-11-03
 ; PRIOR APPLICATION NUMBER: 60/065311
 ; PRIOR FILING DATE: 1997-11-13
 ; PRIOR APPLICATION NUMBER: 60/066364
 ; PRIOR FILING DATE: 1997-11-21
 ; PRIOR APPLICATION NUMBER: 60/077450
 ; PRIOR FILING DATE: 1998-03-10
 ; PRIOR APPLICATION NUMBER: 60/077632
 ; PRIOR FILING DATE: 1998-03-11
 ; PRIOR APPLICATION NUMBER: 60/077641
 ; PRIOR FILING DATE: 1998-03-11
 ; PRIOR APPLICATION NUMBER: 60/077649
 ; PRIOR FILING DATE: 1998-03-11
 ; PRIOR APPLICATION NUMBER: 60/077791
 ; PRIOR FILING DATE: 1998-03-12
 ; PRIOR APPLICATION NUMBER: 60/078004
 ; PRIOR FILING DATE: 1998-03-13
 ; PRIOR APPLICATION NUMBER: 60/078886
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/078936
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/078910
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/078939
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/079294
 ; PRIOR FILING DATE: 1998-03-25
 ; PRIOR APPLICATION NUMBER: 60/079656
 ; PRIOR FILING DATE: 1998-03-26
 ; PRIOR APPLICATION NUMBER: 60/079664
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079689
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079663
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079728
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079786
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079920
 ; PRIOR FILING DATE: 1998-03-30
 ; PRIOR APPLICATION NUMBER: 60/079923
 ; PRIOR FILING DATE: 1998-03-30
 ; PRIOR APPLICATION NUMBER: 60/080105
 ; PRIOR FILING DATE: 1998-03-31
 ; PRIOR APPLICATION NUMBER: 60/080107
 ; PRIOR FILING DATE: 1998-03-31
 ; PRIOR APPLICATION NUMBER: 60/080165
 ; PRIOR FILING DATE: 1998-03-31
 ; PRIOR APPLICATION NUMBER: 60/080194
 ; PRIOR FILING DATE: 1998-03-31
 ; PRIOR APPLICATION NUMBER: 60/080327
 ; PRIOR FILING DATE: 1998-04-01
 ; PRIOR APPLICATION NUMBER: 60/080328
 ; PRIOR FILING DATE: 1998-04-01
 ; PRIOR APPLICATION NUMBER: 60/080333
 ; PRIOR FILING DATE: 1998-04-01
 ; PRIOR APPLICATION NUMBER: 60/080334
 ; PRIOR FILING DATE: 1998-04-01
 ; PRIOR APPLICATION NUMBER: 60/081070
 ; PRIOR FILING DATE: 1998-04-08
 ; PRIOR APPLICATION NUMBER: 60/081049
 ; PRIOR FILING DATE: 1998-04-08
 ; PRIOR APPLICATION NUMBER: 60/081071
 ; PRIOR FILING DATE: 1998-04-08
 ; PRIOR APPLICATION NUMBER: 60/081195

1 PRIOR FILING DATE: 1998-04-08
2 PRIOR APPLICATION NUMBER: 60/081203
3 PRIOR FILING DATE: 1998-04-09
4 PRIOR APPLICATION NUMBER: 60/081229
5 PRIOR FILING DATE: 1998-04-09
6 PRIOR APPLICATION NUMBER: 60/081955
7 PRIOR FILING DATE: 1998-04-15
8 PRIOR APPLICATION NUMBER: 60/081817
9 PRIOR FILING DATE: 1998-04-15
10 PRIOR APPLICATION NUMBER: 60/081819
11 PRIOR FILING DATE: 1998-04-15
12 PRIOR APPLICATION NUMBER: 60/081952
13 PRIOR FILING DATE: 1998-04-15
14 PRIOR APPLICATION NUMBER: 60/081838
15 PRIOR FILING DATE: 1998-04-15
16 PRIOR APPLICATION NUMBER: 60/082568
17 PRIOR FILING DATE: 1998-04-21
18 PRIOR APPLICATION NUMBER: 60/082569
19 PRIOR FILING DATE: 1998-04-21
20 PRIOR APPLICATION NUMBER: 60/082704
21 PRIOR FILING DATE: 1998-04-22
22 PRIOR APPLICATION NUMBER: 60/082796
23 PRIOR FILING DATE: 1998-04-23
24 PRIOR APPLICATION NUMBER: 60/083336
25 PRIOR FILING DATE: 1998-04-27
26 PRIOR APPLICATION NUMBER: 60/083322
27 PRIOR FILING DATE: 1998-04-28
28 PRIOR APPLICATION NUMBER: 60/083392
29 PRIOR FILING DATE: 1998-04-29
30 PRIOR APPLICATION NUMBER: 60/083495
31 PRIOR FILING DATE: 1998-04-29
32 PRIOR APPLICATION NUMBER: 60/083496
33 PRIOR FILING DATE: 1998-04-29
34 PRIOR APPLICATION NUMBER: 60/083499
35 PRIOR FILING DATE: 1998-04-29
36 PRIOR APPLICATION NUMBER: 60/083545
37 PRIOR FILING DATE: 1998-04-29
38 PRIOR APPLICATION NUMBER: 60/083554
39 PRIOR FILING DATE: 1998-04-29
40 PRIOR APPLICATION NUMBER: 60/083558
41 PRIOR FILING DATE: 1998-04-29
42 PRIOR APPLICATION NUMBER: 60/083559
43 PRIOR FILING DATE: 1998-04-29
44 PRIOR APPLICATION NUMBER: 60/083500
45 PRIOR FILING DATE: 1998-04-29
46 PRIOR APPLICATION NUMBER: 60/083742
47 PRIOR FILING DATE: 1998-04-30
48 PRIOR APPLICATION NUMBER: 60/084366
49 PRIOR FILING DATE: 1998-05-05
50 PRIOR APPLICATION NUMBER: 60/084414
51 PRIOR FILING DATE: 1998-05-06
52 PRIOR APPLICATION NUMBER: 60/084441
53 PRIOR FILING DATE: 1998-05-06
54 PRIOR APPLICATION NUMBER: 60/084637
55 PRIOR FILING DATE: 1998-05-07
56 PRIOR APPLICATION NUMBER: 60/084639
57 PRIOR FILING DATE: 1998-05-07
58 PRIOR APPLICATION NUMBER: 60/084640
59 PRIOR FILING DATE: 1998-05-07
60 PRIOR APPLICATION NUMBER: 60/084598
61 PRIOR FILING DATE: 1998-05-07
62 PRIOR APPLICATION NUMBER: 60/084600
63 PRIOR FILING DATE: 1998-5-07
64 PRIOR APPLICATION NUMBER: 60/084627
65 PRIOR FILING DATE: 1998-05-07
66 PRIOR APPLICATION NUMBER: 60/084643
67 PRIOR FILING DATE: 1998-05-07

1 PRIOR APPLICATION NUMBER: 60/085339
2 PRIOR FILING DATE: 1998-05-13
3 PRIOR APPLICATION NUMBER: 60/085338
4 PRIOR FILING DATE: 1998-05-13
5 PRIOR APPLICATION NUMBER: 60/085323
6 PRIOR FILING DATE: 1998-05-13
7 PRIOR APPLICATION NUMBER: 60/085582
8 PRIOR FILING DATE: 1998-05-15
9 PRIOR APPLICATION NUMBER: 60/085700
10 PRIOR FILING DATE: 1998-05-15
11 PRIOR APPLICATION NUMBER: 60/085689
12 PRIOR FILING DATE: 1998-05-15
13 PRIOR APPLICATION NUMBER: 60/085579
14 PRIOR FILING DATE: 1998-05-15
15 PRIOR APPLICATION NUMBER: 60/085580
16 PRIOR FILING DATE: 1998-05-15
17 PRIOR APPLICATION NUMBER: 60/085573
18 PRIOR FILING DATE: 1998-05-15
19 PRIOR APPLICATION NUMBER: 60/085704
20 PRIOR FILING DATE: 1998-05-15
21 PRIOR APPLICATION NUMBER: 60/085697

Query Match 61.1%; Score 33; DB 9; Length 736;

Best Local Similarity 70.0%; Pred. No. 2.6e+02; Mismatches 3; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMDY 10
Db 331 EPVVVYGMDY 340

Search completed: June 4, 2003, 13:30:39
Job time : 15.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:09:19 ; Search time 9.75 Seconds
(without alignments)
33.195 Million cell updates/sec

Title: AUDET-909-3

Perfect score: 54

Sequence: 1 eevvpxgmdys 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	66.7	341	4	US-08-853-948B-4
2	36	66.7	348	4	US-08-853-948B-5
3	36	66.7	368	4	US-09-697-367-24
4	36	66.7	1068	2	US-08-429-054A-11
5	36	66.7	1068	2	US-07-718-777-7
6	36	66.7	1068	3	US-09-051-341-7
7	34	63.0	140	4	US-08-569-147-76
8	34	63.0	140	4	US-08-569-147-82
9	33	61.1	59	4	US-08-963-851-14
10	33	61.1	378	1	US-08-070-165F-8
11	33	61.1	378	2	US-08-885-418-8
12	33	61.1	801	4	US-09-383-630-6
13	32	59.3	65	6	5177197-51
14	32	59.3	102	2	US-08-580-988A-23
15	32	59.3	152	2	US-08-460-694-4
16	32	59.3	152	3	US-08-744-4
17	32	59.3	152	3	US-07-667-711B-4
18	32	59.3	173	1	US-08-193-977-7
19	32	59.3	189	2	US-08-464-517-21
20	32	59.3	189	2	US-08-246-361A-21
21	32	59.3	189	3	US-08-463-772-21
22	32	59.3	231	5	PCT-US93-05000-21
23	32	59.3	231	3	US-08-926-842B-20
24	32	59.3	236	2	US-08-464-517-22
25	32	59.3	236	2	US-08-246-361A-22
26	32	59.3	236	3	US-08-463-772-22
27	32	59.3	236	5	PCT-US93-05000-22

28	59.3	240	3	US-08-926-842B-21	Sequence 21, Appl
29	59.3	280	3	US-08-464-517-6	Sequence 6, Appl
30	59.3	280	3	US-08-463-772-6	Sequence 6, Appl
31	59.3	289	2	US-08-246-361A-4	Sequence 4, Appl
32	59.3	289	5	PCT-US93-05000-4	Sequence 4, Appl
33	59.3	291	5	PCT-US93-05000-6	Sequence 6, Appl
34	59.3	292	2	US-08-464-517-23	Sequence 23, Appl
35	59.3	292	2	US-08-246-361A-6	Sequence 6, Appl
36	59.3	292	2	US-08-246-361A-23	Sequence 23, Appl
37	59.3	292	3	US-08-463-772-23	Sequence 23, Appl
38	59.3	292	5	PCT-US93-05000-23	Sequence 23, Appl
39	59.3	295	1	US-07-947-120-8	Sequence 8, Appl
40	59.3	295	1	US-08-472-893A-8	Sequence 8, Appl
41	59.3	295	2	US-08-460-694-2	Sequence 2, Appl
42	59.3	295	2	US-08-464-517-19	Sequence 19, Appl
43	59.3	295	2	US-08-464-517-20	Sequence 20, Appl
44	59.3	295	2	US-08-246-361A-19	Sequence 19, Appl
45	59.3	295	2	US-08-246-361A-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-08-853-948B-4
; Sequence 4, Application US/08853948B
; Patent No. 6210943
; GENERAL INFORMATION:
; APPLICANT: AKIHAMA, Toyota
; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE FROM CITRUS AND DNA ENCODING
; FILE REFERENCE: THE SAME
; FILE REFERENCE: 0049-0235-0
; CURRENT APPLICATION NUMBER: US/08/853,948B
; CURRENT FILING DATE: 1997-05-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; TYPE: PRT
; LENGTH: 341
; ORGANISM: Citrus unshiu
; FEATURE:
; OTHER INFORMATION: Xaa at position 109 is one of Ala, Arg, Asn, Asp,
; OTHER INFORMATION: Cys, Gln, Gly, His, Ile, Leu, Lys, Met, Phe,
; OTHER INFORMATION: Pro, Ser, Thr, Trp, Tyr, or Val
US-08-853-948B-4

Query Match 66.7%; Score 36; DB 4; Length 341;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3' VVPXGMDYS 11
|:|:|:|:|:
Db 228 VPPGMDFS 236

RESULT 2
US-08-853-948B-5
; Sequence 5, Application US/08853948B
; Patent No. 6210943
; GENERAL INFORMATION:
; APPLICANT: AKIHAMA, Toyota
; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE FROM CITRUS AND DNA ENCODING
; FILE REFERENCE: THE SAME
; FILE REFERENCE: 0049-0235-0
; CURRENT APPLICATION NUMBER: US/08/853,948B
; CURRENT FILING DATE: 1997-05-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; TYPE: PRT
; LENGTH: 348
; ORGANISM: Citrus unshiu
US-08-853-948B-5

audet-909-3.rai

Wed Jun 4 16:50:37 2003

Query Match 66.7%; Score 36; DB 4; Length 348;
 Best Local Similarity 66.7%; Pred. No. 15; Indels 1; Gaps 0;
 Matches 6; Conservative

Qy 3 VVPXGMDYS 11
 Db 234 VIPPGMDFS 242

RESULT 3
 US-09-697-367-24
 ; Sequence 24, Application US/09697367
 ; Patent No. 6323015
 ; GENERAL INFORMATION:
 ; APPLICANT: Orozco Jr., Emil M.
 ; APPLICANT: Caimi, Perry G.
 ; APPLICANT: Weng, Zude
 ; APPLICANT: Tarczynski, Mitchell
 ; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE
 ; FILE REFERENCE: B1166 US NA
 ; CURRENT APPLICATION NUMBER: US/09/697.367
 ; CURRENT FILING DATE: 2000-10-26
 ; PRIOR APPLICATION NUMBER: 60/084,529
 ; PRIOR FILING DATE: 1998-MAY-07
 ; PRIOR APPLICATION NUMBER: PCT/US99/09865
 ; PRIOR FILING DATE: 1999-MAY-06
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 24
 ; LENGTH: 368
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; US-09-697-367-24

Query Match 66.7%; Score 36; DB 4; Length 368;
 Best Local Similarity 66.7%; Pred. No. 16; Indels 1; Gaps 0;
 Matches 6; Conservative

Qy 3 VVPXGMDYS 11
 Db 217 VIPPGMDFS 225

RESULT 4
 US-08-429-054A-11
 ; Sequence 11, Application US/08429054A
 ; Patent No. 5917126
 ; GENERAL INFORMATION:
 ; APPLICANT: VAN ASSCHE, CHARLES; LANDO, DANIELLE; BRUNEAU,
 ; APPLICANT: JEAN; VOELKER, TONI; GERVAIS, MONICA
 ; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHETASE (SPS),
 ; TITLE OF INVENTION: A PREPARATION METHOD AND CDNA THEREFOR, AND USE OF THE
 ; TITLE OF INVENTION: CDNA FOR MODIFYING SPS EXPRESSION IN PLANT CELLS
 ; NUMBER OF SEQUENCES: 37
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BIERMAN AND MUSERLIAN
 ; STREET: 600 THIRD AVENUE
 ; CITY: NEW YORK
 ; STATE: NEW YORK
 ; COUNTRY: USA
 ; ZIP: 10016
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: FLOPPY DISK
 ; COMPUTER: IBM PC COMPATIBLE
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: WORDPERFECT 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/429,054A
 ; FILING DATE: 26-APR-1995
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 842,337

FILING DATE: 20-March-1992
 APPLICATION NUMBER: PCT/FR 91/00593
 FILING DATE: 18-July-1991
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: French 90402094.9
 FILING DATE: 20-July-1990
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: Charles A. Musserlian
 REGISTRATION NUMBER: 19,683
 REFERENCE/DOCKET NUMBER: 146.1137
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 661-8000
 TELEFAX: (212) 661-8002
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1068
 TYPE: Amino acid
 STRANDEDNESS: Single
 TOPOLOGY: Unknown
 MOLECULE TYPE: Peptide
 US-08-429-054A-11

Query Match 66.7%; Score 36; DB 2; Length 1068;
 Best Local Similarity 66.7%; Pred. No. 53; Indels 1; Gaps 0;
 Matches 6; Conservative

Qy 3 VVPXGMDYS 11
 Db 435 VIPPGMDFS 443

RESULT 5
 US-08-718-777-7
 ; Sequence 7, Application US/08718777
 ; Patent No. 5981852
 ; GENERAL INFORMATION:
 ; APPLICANT: Van Assche, C.
 ; APPLICANT: Lando, D. D.
 ; APPLICANT: Bruneau, J. M.
 ; APPLICANT: Voelker, T.
 ; APPLICANT: Gervais, M.
 ; TITLE OF INVENTION: MODIFICATION OF SUCROSE
 ; TITLE OF INVENTION: PHOSPHATE
 ; TITLE OF INVENTION: SYNTHASE IN PLANTS
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Law Offices of Barbara Rae-Venter
 ; STREET: 260 Sheridan Avenue, Suite 440
 ; CITY: Palo Alto
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94306
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/718,777
 ; FILING DATE: NOT YET ASSIGNED
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/175,471
 ; FILING DATE: 27-DEC-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Barbara Rae-Venter
 ; REGISTRATION NUMBER: 32,750
 ; REFERENCE/DOCKET NUMBER: CGNE.072.02US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415)328-4400
 ; TELEFAX: (415)328-4477

; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1068 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-718-777-7

Query Match 66.7%; Score 36; DB 2; Length 1068;
 Best Local Similarity 66.7%; Pred. No. 53;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
 | : | | | : |
 Db 435 VIPPGMDFS 443

RESULT 6
 US-09-051-341-7
 ; Sequence 7, Application US/09051341
 ; Patent No. 6124528
 ; GENERAL INFORMATION:
 ; APPLICANT: Shewmaker, C. K.
 ; TITLE OF INVENTION: MODIFICATION OF SOLUBLE SOLIDS USING
 ; CURRENT APPLICATION DATA:
 ; FILING DATE: 25-March-1996
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Trujillo, Doreen Yanko
 ; REGISTRATION NUMBER: 35,719
 ; REFERENCE/DOCKET NUMBER: CARP-0047
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (215) 568-3100
 ; TELEFAX: (215) 568-3439
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 140 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-051-341-7

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/051.341
 FILING DATE:
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US96/17351
 FILING DATE: 25-OCT-1996
 APPLICATION NUMBER: US 08/549,016
 FILING DATE: 27-OCT-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/372,200
 FILING DATE: 12-JAN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Barbara Rae-Venter, Ph.D.,
 REGISTRATION NUMBER: 32,750
 REFERENCE/DOCKET NUMBER: CGNE.110.02US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415)328-4400
 TELEFAX: (415)328-4477

INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1068 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-051-341-7

Query Match 66.7%; Score 36; DB 3; Length 1068;
 Best Local Similarity 66.7%; Pred. No. 53;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
 | : | | | : |
 Db 435 VIPPGMDFS 443

RESULT 7
 US-08-569-147-76
 ; Sequence 76, Application US/08569147
 ; Patent No. 6180377
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; TITLE OF INVENTION: HUMANISED ANTIBODIES
 ; NUMBER OF SEQUENCES: 95
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
 ; ADDRESS: No. 6180377ris, LLP
 ; STREET: One Liberty Place - 46th Floor
 ; CITY: Philadelphia
 ; STATE: PA
 ; COUNTRY: U.S.A.
 ; ZIP: 19103
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/569,147
 ; FILING DATE: 25-March-1996
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Trujillo, Doreen Yanko
 ; REGISTRATION NUMBER: 35,719
 ; REFERENCE/DOCKET NUMBER: CARP-0047
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (215) 568-3100
 ; TELEFAX: (215) 568-3439
 ; INFORMATION FOR SEQ ID NO: 76:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 140 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-569-147-76

Query Match 63.0%; Score 34; DB 4; Length 140;
 Best Local Similarity 75.0%; Pred. No. 13;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMDY 10
 | | | | |
 Db 122 VVPTGFDY 129

RESULT 8
 US-08-569-147-82
 ; Sequence 82, Application US/08569147
 ; Patent No. 6180377
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; TITLE OF INVENTION: HUMANISED ANTIBODIES
 ; NUMBER OF SEQUENCES: 95
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
 ; ADDRESS: No. 6180377ris, LLP
 ; STREET: One Liberty Place - 46th Floor
 ; CITY: Philadelphia
 ; STATE: PA
 ; COUNTRY: U.S.A.
 ; ZIP: 19103
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/569,147

FILING DATE: 25-March-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Trujillo, Doreen Yanko
REGISTRATION NUMBER: 35,719
REFERENCE/DOCKET NUMBER: CARP-0047
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 82:
SEQUENCE CHARACTERISTICS:
LENGTH: 140 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-569-147-82

Query Match 63.0%; Score 34; DB 4; Length 140;
Best Local Similarity 75.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 VVPXGMDY 10
Db 122 VVPTGFDY 129

RESULT 9
US-08-963-851-14
Sequence 14, Application US/08963851
Patent No. 6300116
GENERAL INFORMATION:
APPLICANT: VAN DER OSTEN, CLAUS
APPLICANT: HALKIER, TORDEN
APPLICANT: ANDERSEN, CARSTEN
APPLICANT: BAUDITZ, PETER
APPLICANT: HANSEN, PETER KAMP
TITLE OF INVENTION: PROTEASE VARIANTS AND COMPOSITIONS
FILE REFERENCE: 4946,200-US
CURRENT APPLICATION NUMBER: US/08/963,851
CURRENT FILING DATE: 1997-11-04
NUMBER OF SEQ ID NOS: 35
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 14
LENGTH: 59
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-08-963-851-14

Query Match 61.1%; Score 33; DB 4; Length 59;
Best Local Similarity 45.5%; Pred. No. 7.5;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 BEVVPXGMDYS 11
Db 38 EKHPGGLEYS 48

RESULT 10
US-08-070-165F-8
Sequence 8, Application US/08070165F
Patent No. 5750365
GENERAL INFORMATION:
APPLICANT: Chiu, Ing-Ming
APPLICANT: Poulin, Matthew L
TITLE OF INVENTION: Acidic Fibroblast Growth Factor (aFGF)
CORRESPONDENCE ADDRESS:
NUMBER OF SEQUENCES: 12
ADDRESSEE: Ing-Ming Chiu
STREET: S2052 Davis Medical Research Center, 480 West
CITY: Columbus
STATE: Ohio
COUNTRY: USA

ZIP: 43210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/070,165F
FILING DATE:
CLASSIFICATION: 435
TELECOMMUNICATION INFORMATION:
TELEPHONE: (614)-293-8093
TELEFAX: (614)-293-5631
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 378 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-070-165F-8

Query Match 61.1%; Score 33; DB 1; Length 378;
Best Local Similarity 85.7%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 PXGMDYS 11
Db 139 PPGMDYS 145

RESULT 11
US-08-885-418-8
Sequence 8, Application US/08885418
Patent No. 5925528
GENERAL INFORMATION:
APPLICANT: Chiu, Ing-Ming
APPLICANT: Poulin, Matthew L
TITLE OF INVENTION: Acidic Fibroblast Growth Factor (aFGF)
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ing-Ming Chiu
STREET: S2052 Davis Medical Research Center, 480 West
CITY: Columbus
STATE: Ohio
COUNTRY: USA
ZIP: 43210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/885,418
FILING DATE:
CLASSIFICATION: 435
TELECOMMUNICATION INFORMATION:
TELEPHONE: (614)-293-8093
TELEFAX: (614)-293-5631
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 378 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-885-418-8

Query Match 61.1%; Score 33; DB 2; Length 378;
Best Local Similarity 85.7%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 PXGMDYS 11
Db 139 PPGMDYS 145

Db 139 PPGMDYS 145

RESULT 12

US-09-383-630-6

; Sequence 6, Application US/09383630A

; Patent No. 6265632

; GENERAL INFORMATION:

; APPLICANT: Amer Yayon et al.

; TITLE OF INVENTION: ANIMAL MODEL FOR FIBROBLAST GROWTH

; FACTOR RECEPTOR ASSOCIATED

; CHONDRODYSPLASIA

; NUMBER OF SEQUENCES: 18

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Mark M. Friedman c/o Anthony Castorina

; STREET: 2001 Jefferson Davis Highway, Suite 207

; CITY: Arlington

; STATE: Virginia

; COUNTRY: United States of America

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk

; COMPUTER: Twinhead* Slimnote-890TX

; OPERATING SYSTEM: MS DOS version 6.2,

; Windows version 3.11

; SOFTWARE: Word for Windows version 2.0 converted

; to an ASCII file

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/383.630A

; FILING DATE: 26-Aug-1999

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: <Unknown>

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Friedman, Mark M.

; REGISTRATION NUMBER: 33,883

; REFERENCE/DOCKET NUMBER: 1402/2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 972-3-5625553

; TELEFAX: 972-3-5625554

; TELEX: <Unknown>

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 801

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-09-383-630-6

Query Match 61.1%; Score 33; DB 4; Length 801;

Best Local Similarity 85.7%; Pred. No. 1.5e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PPGMDYS 11

Db 566 PPGMDYS 572

RESULT 13

5177197-51

; Patent No. 5177197

; APPLICANT: KANZAKI, TETSUO; OLOFSSON, ANDERS; MOREN, ANITA;

; WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAESSON-WELSH,

; LENA; HELDIN, CARL-HENRIK

; TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING

; HUMAN TRANSFORMING GROWTH FACTOR-BETA1-BINDING PROTEIN

; NUMBER OF SEQUENCES: 53

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/487,343

; FILING DATE: 27-FEB-1990

; SEQ ID NO: 51:

5177197-51

LENGTH: 65

Query Match

Best Local Similarity 59.3%; Score 32; DB 6; Length 65;

Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 BEVVPXGMDYS 11

Db 52 KEICPGMGYT 62

RESULT 14

US-08-580-988A-23

; Sequence 23, Application US/08580988A

; Patent No. 5856161

; GENERAL INFORMATION:

; APPLICANT: Aggarwal et al.

; TITLE OF INVENTION: Tumor Necrosis Factor

; TITLE OF INVENTION: Receptor-I-Associated Protein Kinase And Methods

; TITLE OF INVENTION: For Its Use

; NUMBER OF SEQUENCES: 27

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dr. Benjamin A. Adler

; STREET: 8011 Candle Lane

; CITY: Houston

; STATE: Texas

; COUNTRY: USA

; ZIP: 77071

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 1.44 Mb floppy disk

; COMPUTER: Apple Macintosh

; OPERATING SYSTEM: Macintosh

; SOFTWARE: Microsoft Word for Macintosh

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/580,988A

; FILING DATE: January 3, 1996

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Benjamin Aaron Adler, Ph.D., J.D.

; REGISTRATION NUMBER: 35,423

; REFERENCE/DOCKET NUMBER: D5721CIP2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 713-777-2321

; TELEFAX: 713-777-6908

; INFORMATION FOR SEQ ID NO: 23:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 102 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE:

; DESCRIPTION: protein

; HYPOTHETICAL: no

; ANTI-SENSE: no

; FRAGMENT TYPE: internal

; ORIGINAL SOURCE:

US-08-580-988A-23

Query Match 59.3%; Score 32; DB 2; Length 102;

Best Local Similarity 60.0%; Pred. No. 22;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 BEVVPXGMDY 10

Db 24 BEVFPPLAMNY 33

RESULT 15

US-08-460-694-4

```

; Sequence 4, Application US/08460694
; Patent No. 5858655
; GENERAL INFORMATION:
; APPLICANT: Arnold, Andrew
; TITLE OF INVENTION: PRADI Cyclin and its cDNA
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,694
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McConathy, Evelyn H.
; REGISTRATION NUMBER: 35,279
; REFERENCE/DOCKET NUMBER: 0609.4070002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 152 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
; US-08-460-694-4

```

```

Query Match          59.3%; Score 32; DB 2; Length 152;
Best Local Similarity 60.0%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      1 BEVVPXGMDY 10
        ||| | |
Db      20 BEVFPPLAMNY 29

```

```

Search completed: June 4, 2003, 13:14:59
Job time : 9.75 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:04:09 ; Search time 28.25 Seconds
(without alignments)
51.885 Million cell updates/sec

Title: AUDET-909-4 SEQ 47
Perfect score: 55
Sequence: 1 (evvpvmsys 11-2) ALL COMMON AR 3

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq 101002:*

1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	50	90.9	11	23	Hepatitis C virus
2	50	90.9	11	23	Hepatitis C virus
3	50	90.9	11	23	Hepatitis C virus
4	50	90.9	11	23	Hepatitis C virus
5	50	90.9	11	23	Hepatitis C virus
6	50	90.9	11	23	Hepatitis C virus
7	50	90.9	11	23	Hepatitis C virus
8	50	90.9	11	23	Hepatitis C virus
9	50	90.9	11	23	Hepatitis C virus
10	50	90.9	11	23	Hepatitis C virus

11	50	90.9	11	23	ABB80568	Hepatitis C virus
12	46	83.6	11	23	ABB80524	Hepatitis C virus
13	46	83.6	11	23	ABB80528	Hepatitis C virus
14	46	83.6	11	23	ABB80529	Hepatitis C virus
15	46	83.6	11	23	ABB80561	Hepatitis C virus
16	46	83.6	11	23	ABB80562	Hepatitis C virus
17	45	81.8	11	23	ABB80527	Hepatitis C virus
18	45	81.8	11	23	ABB80527	Hepatitis C virus
19	45	81.8	11	23	ABB80535	Hepatitis C virus
20	45	81.8	11	23	ABB80536	Hepatitis C virus
21	45	81.8	11	23	ABB80539	Hepatitis C virus
22	45	81.8	11	23	ABB80540	Hepatitis C virus
23	45	81.8	11	23	ABB80558	Hepatitis C virus
24	44	80.0	11	23	ABB80560	Hepatitis C virus
25	44	80.0	11	23	ABB80544	Hepatitis C virus
26	44	80.0	11	23	ABB80545	Hepatitis C virus
27	44	80.0	11	23	ABB80549	Hepatitis C virus
28	44	80.0	11	23	ABB80552	Hepatitis C virus
29	44	80.0	11	23	ABB80553	Hepatitis C virus
30	42	76.4	11	23	ABB80530	Hepatitis C virus
31	41	74.5	11	23	ABB80538	Hepatitis C virus
32	41	74.5	11	23	ABB80542	Hepatitis C virus
33	41	74.5	11	23	ABB80543	Hepatitis C virus
34	41	74.5	20	20	AAU76810	Hepatitis C virus
35	40	72.7	11	23	ABB80537	Hepatitis C virus
36	40	72.7	11	23	ABB80541	Hepatitis C virus
37	40	72.7	11	23	ABB80547	Hepatitis C virus
38	40	72.7	11	23	ABB80548	Hepatitis C virus
39	40	72.7	11	23	ABB80551	Hepatitis C virus
40	40	72.7	11	23	ABB80556	Hepatitis C virus
41	40	72.7	11	23	ABB80557	Hepatitis C virus
42	40	72.7	11	23	ABB80562	Novel human diagno
43	40	72.7	1022	22	ABG03621	Novel human diagno
44	40	72.7	1022	22	ABG05826	Novel human diagno
45	39	70.9	11	23	ABB80546	Hepatitis C virus

ALIGNMENTS

RESULT 1

ABB80521
ID ABB80521 standard; peptide; 11 AA.

XX ABB80521
AC ABB80521

DT 08-OCT-2002 (first entry)

DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #1.

KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide; virucide.

OS Synthetic.

XX Key
FH Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with residue 7"

FT Modified-site 11 /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease -
 XX Claim 17; Page 64; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX Sequence 11 AA;
 SQ Query Match 90.9%; Score 50; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.0033;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 EEVVPVGMYS 11
 Db 1 EEVVPXGMSYS 11
 ||||| |||||
 RESULT 2
 ABB80522
 ID ABB80522 standard; peptide; 11 AA.
 AC ABB80522;
 XX 08-OCT-2002 (first entry)
 DT Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #2.
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 DE virucide.
 KW Synthetic.
 KW Key Location/Qualifiers
 XX Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 9 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 XX 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease -
 XX Claim 17; Page 64; 69pp; English.

XX Claim 17; Page 64; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX Sequence 11 AA;
 SQ Query Match 90.9%; Score 50; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.0033;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 EEVVPVGMYS 11
 Db 1 EEVVPXGMSYS 11
 ||||| |||||
 RESULT 3
 ABB80525
 ID ABB80525 standard; peptide; 11 AA.
 AC ABB80525;
 XX 08-OCT-2002 (first entry)
 DT Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #5.
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 DE virucide.
 KW Synthetic.
 KW Key Location/Qualifiers
 XX Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
 FT residue 7"
 FT Misc-difference 8 /note= "D-form residue"
 FT Modified-site 11 /note= "C-terminal amide"
 FT WO200208251-A2.
 XX 31-JAN-2002.
 XX 19-JUL-2001; 2001WO-US23169.
 XX 21-JUL-2000; 2000US-220101P.
 XX (CORV-) CORVAS INT INC.
 XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.
 XX Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease -
 XX Claim 17; Page 64; 69pp; English.
 XX The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A

CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 90.9%; Score 50; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.0033;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPVGMYS 11

Db 1 EEVVPVGMYS 11

RESULT 4

ABB80526

ID ABB80526 standard; peptide; 11 AA.

XX

AC ABB80526;

XX

DT 08-OCT-2002 (first entry)

XX

DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #6.

XX

KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide; virucide.

XX

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Modified-site 1

FT /note= "N-terminal acetyl"

FT Modified-site 6

FT /note= "Norvalyl carbonyl forming keto-amide linkage with residue 7"

FT Misc-difference 8

FT /note= "D-form residue"

FT Misc-difference 9

FT /note= "D-form residue"

FT Modified-site 11

FT /note= "C-terminal amide"

XX

PN WO200208251-A2.

XX

PD 31-JAN-2002.

XX

PF 19-JUL-2001; 2001WO-US23169.

XX

PR 21-JUL-2000; 2000US-220101P.

XX

PA (CORV-) CORVAS INT INC.

XX

PI Lim-wilby M, Levy OE, Brunck TK;

XX

DR WPI; 2002-361643/39.

XX

PT Novel peptide compound having hepatitis C virus protease inhibitory activity useful for treating disorders associated with hepatitis C virus protease

XX

PS Claim 17; Page 64; 69pp; English.

XX

CC The sequence represents a peptide compound of the invention having hepatitis C virus (HCV) protease inhibitory activity. The peptides of the invention are alpha-ketoamide peptide analogues. The peptides have virucide activity, and are useful for treating and in the manufacture of a medicament to treat disorders associated with HCV protease. A pharmaceutical composition comprising the peptide as an active ingredient is useful for treating disorders associated with hepatitis C virus.

CC

XX

SQ Sequence 11 AA;

Query Match

90.9%; Score 50; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.0033;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPVGMYS 11

Db 1 EEVVPVGMYS 11

RESULT 5

ABB80559

ID ABB80559 standard; peptide; 11 AA.

XX

AC ABB80559;

XX

DT 08-OCT-2002 (first entry)

XX

DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #39.

XX

KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide; virucide.

XX

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Modified-site 1

FT /note= "N-terminal acetyl"

FT Modified-site 6

FT /note= "Norvalyl carbonyl forming keto-amide linkage with residue 7"

FT Misc-difference 8

FT /note= "D-form residue"

FT Modified-site 8

FT /note= "Oxymethionine"

FT Modified-site 11

FT /note= "C-terminal amide"

XX

PN WO200208251-A2.

XX

PD 31-JAN-2002.

XX

PF 19-JUL-2001; 2001WO-US23169.

XX

PR 21-JUL-2000; 2000US-220101P.

XX

PA (CORV-) CORVAS INT INC.

XX

PI Lim-wilby M, Levy OE, Brunck TK;

XX

DR WPI; 2002-361643/39.

XX

PT Novel peptide compound having hepatitis C virus protease inhibitory activity useful for treating disorders associated with hepatitis C virus protease

XX

PS Claim 17; Page 65; 69pp; English.

XX

CC The sequence represents a peptide compound of the invention having hepatitis C virus (HCV) protease inhibitory activity. The peptides of the invention are alpha-ketoamide peptide analogues. The peptides have virucide activity, and are useful for treating and in the manufacture of a medicament to treat disorders associated with HCV protease. A pharmaceutical composition comprising the peptide as an active ingredient is useful for treating disorders associated with hepatitis C virus.

CC

XX

SQ Sequence 11 AA;

Query Match

90.9%; Score 50; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.0033;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPVGMYS 11

Db 1 EEVVPVGMYS 11

FT Modified-site 6 /note= "Norleucyl carbonyl forming keto-amide linkage
FT Modified-site 11 with residue 7"
FT Modified-site 11 /note= "C-terminal amide"
XX WO200208251-A2.
XX 31-JAN-2002.
XX
XX 19-JUL-2001; 2001WO-US23169.
XX 21-JUL-2000; 2000US-220101P.
XX (CORV-) CORVAS INT INC.
XX Lim-wilby M, Levy OE, Brunck TK;
XX WPI; 2002-361643/39.
XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease -
XX
XX Claim 17; Page 65; 69pp; English.
XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX
XX Sequence 11 AA;
XX
XX Query Match 90.9%; Score 50; DB 23; Length 11;
XX Best Local Similarity 90.9%; Pred. No. 0.0033;
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPVGMYSYS 11
DB 1 EEVVPXGMYSYS 11
RESULT 9
ABB80566
ID ABB80566 standard; peptide; 11 AA.
XX ABB80566;
XX 08-OCT-2002 (first entry)
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #46.
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX virucide.
XX Synthetic.
XX Key Location/Qualifiers
XX Modified-site 1 /note= "N-terminal acetyl"
XX Modified-site 6 /note= "2-aminoisobutyryl carbonyl residue forming a
XX keto-amide linkage with residue 7"
XX Modified-site 11 /note= "C-terminal amide"
XX WO200208251-A2.
XX 31-JAN-2002.
XX
XX Modified-site 6 /note= "N-terminal acetyl"
XX Modified-site 6 /note= "2-aminoisobutyryl carbonyl residue forming a
XX keto-amide linkage with residue 7"
XX Modified-site 11 /note= "C-terminal amide"
XX WO200208251-A2.
XX 31-JAN-2002.
XX

PF 19-JUL-2001; 2001WO-US23169.
XX 21-JUL-2000; 2000US-220101P.
XX (CORV-) CORVAS INT INC.
XX Lim-wilby M, Levy OE, Brunck TK;
XX WPI; 2002-361643/39.
XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease -
XX
XX Claim 17; Page 65; 69pp; English.
XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX
XX Sequence 11 AA;
XX
XX Query Match 90.9%; Score 50; DB 23; Length 11;
XX Best Local Similarity 90.9%; Pred. No. 0.0033;
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPVGMYSYS 11
DB 1 EEVVPXGMYSYS 11
RESULT 10
ABB80567
ID ABB80567 standard; peptide; 11 AA.
XX ABB80567;
XX 08-OCT-2002 (first entry)
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #47.
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX virucide.
XX Synthetic.
XX Key Location/Qualifiers
XX Modified-site 1 /note= "N-terminal acetyl"
XX Modified-site 6 /note= "(s,s)allothreonyl carbonyl residue forming a
XX keto-amide linkage with residue 7"
XX Modified-site 11 /note= "C-terminal amide"
XX WO200208251-A2.
XX 31-JAN-2002.
XX
XX Modified-site 6 /note= "N-terminal acetyl"
XX Modified-site 6 /note= "2-aminoisobutyryl carbonyl residue forming a
XX keto-amide linkage with residue 7"
XX Modified-site 11 /note= "C-terminal amide"
XX WO200208251-A2.
XX 31-JAN-2002.
XX
XX Modified-site 6 /note= "N-terminal acetyl"
XX Modified-site 6 /note= "2-aminoisobutyryl carbonyl residue forming a
XX keto-amide linkage with residue 7"
XX Modified-site 11 /note= "C-terminal amide"
XX WO200208251-A2.
XX 31-JAN-2002.
XX

PT Novel peptide compound having hepatitis C virus protease inhibitory
 PT activity useful for treating disorders associated with hepatitis C
 PT virus protease
 PS
 PS Claim 17; Page 65; 69pp; English.
 XX
 CC The sequence represents a peptide compound of the invention having
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 CC invention are alpha-ketoamide peptide analogues. The peptides have
 CC virucide activity, and are useful for treating and in the manufacture of
 CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 SQ Sequence 11 AA;

Query Match 90.9%; Score 50; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.0033;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPVGMYS 11
 ||||| |||||
 DB 1 EEVVPXGMYS 11

RESULT 11
 ABB80568
 ID ABB80568 standard; peptide; 11 AA.

XX ABB80568;
 XX
 XX 08-OCT-2002 (first entry)
 XX
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #48.
 XX
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 XX virucide.
 XX Synthetic.

XX Key Location/Qualifiers
 XX Modified-site 1 /note= "N-terminal acetyl"
 XX Modified-site 6
 XX Modified-site 11 /note= "alpha-propionyl-glycyl-carbonyl residue forming
 XX a keto-amide linkage with residue 7"
 XX Modified-site 11 /note= "C-terminal amide"
 XX
 XX WO200208251-A2.

XX 31-JAN-2002.
 XX
 XX 19-JUL-2001; 2001WO-US23169.
 XX
 XX 21-JUL-2000; 2000US-220101P.
 XX
 XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory
 XX activity useful for treating disorders associated with hepatitis C
 XX virus protease

XX Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having
 XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 XX invention are alpha-ketoamide peptide analogues. The peptides have
 XX virucide activity, and are useful for treating and in the manufacture of

CC a medicament to treat disorders associated with HCV protease. A
 CC pharmaceutical composition comprising the peptide as an active ingredient
 CC is useful for treating disorders associated with hepatitis C virus.
 XX
 SQ Sequence 11 AA;

Query Match 90.9%; Score 50; DB 23; Length 11;
 Best Local Similarity 90.9%; Pred. No. 0.0033;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EEVVPVGMYS 11
 ||||| |||||
 DB 1 EEVVPXGMYS 11

RESULT 12
 ABB80524
 ID ABB80524 standard; peptide; 11 AA.

XX ABB80524;
 XX
 XX 08-OCT-2002 (first entry)
 XX
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #4.
 XX
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
 XX virucide.
 XX Synthetic.

XX Key Location/Qualifiers
 XX Modified-site 1 /note= "N-terminal acetyl"
 XX Modified-site 6
 XX Modified-site 11 /note= "Norvalyl carbonyl forming keto-amide linkage with
 XX residue 7"
 XX Misc-difference 9 /note= "D-form residue"
 XX Modified-site 11 /note= "C-terminal amide"
 XX
 XX WO200208251-A2.

XX 31-JAN-2002.
 XX
 XX 19-JUL-2001; 2001WO-US23169.
 XX
 XX 21-JUL-2000; 2000US-220101P.
 XX
 XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;
 XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory
 XX activity useful for treating disorders associated with hepatitis C
 XX virus protease

XX Claim 17; Page 64; 69pp; English.

XX The sequence represents a peptide compound of the invention having
 XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
 XX invention are alpha-ketoamide peptide analogues. The peptides have
 XX virucide activity, and are useful for treating and in the manufacture of
 XX a medicament to treat disorders associated with HCV protease. A
 XX pharmaceutical composition comprising the peptide as an active ingredient
 XX is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 83.6%; Score 46; DB 23; Length 11;
 Best Local Similarity 81.8%; Pred. No. 0.02;

Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPVGMYSYS 11
DB 1 EEVVPXGMDYS 11

RESULT 13
ABB80528
ID ABB80528 standard; peptide; 11 AA.
XX
AC ABB80528;
XX
DT 08-OCT-2002 (first entry)
XX
DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #9.
XX
KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
virucide.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
residue 7"
FT Misc-difference 8 /note= "D-form residue"
FT Misc-difference 9 /note= "D-form residue"
FT Modified-site 11 /note= "C-terminal amide"
FT
FT
PN WO200208251-A2.
XX
PD 31-JAN-2002.
XX
PF 19-JUL-2001; 2001WO-US23169.
XX
PR 21-JUL-2000; 2000US-220101P.
XX
PA (CORV-) CORVAS INT INC.
XX
PI Lim-wilby M, Levy OE, Brunck TK;
XX
DR WPI; 2002-361643/39.
XX
PT Novel peptide compound having hepatitis C virus protease inhibitory
activity useful for treating disorders associated with hepatitis C
virus protease
XX
PS Claim 17; Page 64; 69pp; English.
XX
CC The sequence represents a peptide compound of the invention having
hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
invention are alpha-ketoamide peptide analogues. The peptides have
virucide activity, and are useful for treating and in the manufacture of
a medicament to treat disorders associated with HCV protease. A
pharmaceutical composition comprising the peptide as an active ingredient
is useful for treating disorders associated with hepatitis C virus.
XX
SQ Sequence 11 AA;
Query Match 83.6%; Score 46; DB 23; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.02; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPVGMYSYS 11
DB 1 EEVVPXGMDYS 11

RESULT 14
ABB80561
ID ABB80561 standard; peptide; 11 AA.
XX
AC ABB80561;
XX
DT 08-OCT-2002 (first entry)

ABB80529
ID ABB80529 standard; peptide; 11 AA.
XX
AC ABB80529;
XX
DT 08-OCT-2002 (first entry)
XX
DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #9.
XX
KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
virucide.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
residue 7"
FT Misc-difference 8 /note= "D-form residue"
FT Misc-difference 9 /note= "D-form residue"
FT Modified-site 11 /note= "C-terminal amide"
FT
FT
PN WO200208251-A2.
XX
PD 31-JAN-2002.
XX
PF 19-JUL-2001; 2001WO-US23169.
XX
PR 21-JUL-2000; 2000US-220101P.
XX
PA (CORV-) CORVAS INT INC.
XX
PI Lim-wilby M, Levy OE, Brunck TK;
XX
DR WPI; 2002-361643/39.
XX
PT Novel peptide compound having hepatitis C virus protease inhibitory
activity useful for treating disorders associated with hepatitis C
virus protease
XX
PS Claim 17; Page 64; 69pp; English.
XX
CC The sequence represents a peptide compound of the invention having
hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
invention are alpha-ketoamide peptide analogues. The peptides have
virucide activity, and are useful for treating and in the manufacture of
a medicament to treat disorders associated with HCV protease. A
pharmaceutical composition comprising the peptide as an active ingredient
is useful for treating disorders associated with hepatitis C virus.
XX
SQ Sequence 11 AA;
Query Match 83.6%; Score 46; DB 23; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.02; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPVGMYSYS 11
DB 1 EEVVPXGMDYS 11

RESULT 15
ABB80561
ID ABB80561 standard; peptide; 11 AA.
XX
AC ABB80561;
XX
DT 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #41.
DE
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX
XX Synthetic.
OS
XX
XX Location/Qualifiers
FH Key
FT Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"
FT Misc-difference 8 /note= "D-form residue"
FT Modified-site 8 /note= "Oxymethionine"
FT Modified-site 11
FT /note= "C-terminal amide"
XX
XX WO200208251-A2.
XX
XX
XX 31-JAN-2002.
XX
XX 19-JUL-2001; 2001WO-US23169.
XX
XX 21-JUL-2000; 2000US-220101P.
XX
XX (CORV-) CORVAS INT INC.
XX
XX Lim-wilby M, Levy OE, Brunck TK;
XX
XX WPI; 2002-361643/39.
XX
XX Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease -
XX
XX Claim 17; Page 65; 69pp; English.
XX
XX The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX
SQ Sequence 11 AA;

Query Match 83.6%; Score 46; DB 23; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.02;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPVGMYS 11
Db 1 EEVVPXGMDYS 11
|||||
|||||

Search completed: June 4, 2003, 13:11:13
Job time : 28.25 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:07:09 ; Search time 22.5 Seconds
(without alignments)
100.734 Million cell updates/sec

Title: AUDET-909-4

Perfect score: 55

Sequence: 1 eevpvgmsys 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	72.7	583	5 Q9BHA5	Q9bha5 plasmodium
2	40	72.7	583	5 Q9BH83	Q9bh83 plasmodium
3	39	70.9	219	17 Q971S2	Q971s2 sulfolobus
4	39	70.9	541	16 Q98BP5	Q98bp5 rhizobium 1
5	38	69.1	209	16 Q9RE56	Q9re56 fusobacteri
6	37	67.3	84	16 Q97DE7	Q97de7 clostridium
7	37	67.3	584	16 Q9R8K6	Q9r8k6 thermoanaer
8	37	67.3	933	5 Q9S339	Q9s339 encephalito
9	37	67.3	1305	5 Q9V7C7	Q9v7c7 drosophila
10	36	65.5	156	3 Q12479	Q12479 saccharomyc
11	36	65.5	234	2 Q32330	Q32330 clostridium
12	36	65.5	273	2 Q9WVW3	Q9wvw3 mycobacteri
13	36	65.5	290	4 Q96MU1	Q96mu1 homo sapien
14	36	65.5	387	16 Q98FX1	Q98fx1 rhizobium 1
15	36	65.5	471	11 Q8R126	Q8r126 mus musculu
16	36	65.5	484	11 Q8VD18	Q8vd18 mus musculu

17	36	65.5	573	16 Q8UH45	Q8uh45 agrobacteri
18	36	65.5	664	16 Q9X104	Q9x104 thermotoga
19	36	65.5	827	11 Q8VCT2	Q8vct2 mus musculu
20	36	65.5	840	3 Q9URY8	Q9ury8 schizosacch
21	36	65.5	1063	16 Q8EG86	Q8eg86 fusobacteri
22	36	65.5	3472	1 O74056	O74056 cenarchaeum
23	35	63.6	227	16 Q9RZU8	Q9rzu8 deinococcus
24	35	63.6	556	4 O43733	O43733 homo sapien
25	35	63.6	600	5 Q9W2C1	Q9w2c1 drosophila
26	35	63.6	630	2 O87110	O87110 comamonas a
27	35	63.6	670	11 Q01487	Q01487 rattus norv
28	35	63.6	697	5 Q9N632	Q9n632 leishmania
29	35	63.6	955	5 Q20829	Q20829 caenorhabdi
30	35	63.6	1902	4 Q14122	Q14122 homo sapien
31	34	61.8	103	11 Q9D0H9	Q9d0h9 mus musculu
32	34	61.8	153	13 P79919	P79919 xenopus lae
33	34	61.8	156	11 Q9D8L9	Q9d8l9 mus musculu
34	34	61.8	190	13 O57481	O57481 stizostedio
35	34	61.8	191	11 Q99NB4	Q99nb4 rattus norv
36	34	61.8	240	11 Q9DB09	Q9db09 mus musculu
37	34	61.8	246	2 O05119	O05119 methylobact
38	34	61.8	252	17 O28342	O28342 archaeoglob
39	34	61.8	291	13 Q8QFP4	Q8qfp4 brachydanio
40	34	61.8	317	10 O23220	O23220 arabidopsis
41	34	61.8	326	2 Q9KWR0	Q9kwr0 bacillus sp
42	34	61.8	327	16 O83791	O83791 treponema p
43	34	61.8	388	2 Q8RL92	Q8rl92 pseudomonas
44	34	61.8	420	3 O74949	O74949 schizosacch
45	34	61.8	471	10 Q9FJ24	Q9fj24 arabidopsis

ALIGNMENTS

RESULT 1

Q9BHA5 ID Q9BHA5 PRELIMINARY; PRT; 583 AA.
AC Q9BHA5;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Choline transporter.
GN SCTL.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RA Ben Mamoun C., Gluzman I.Y., Goldberg D.E.;
RT "Plasmodium falciparum choline transporter (pfSCTL) gene."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY007372; AAK14816.1; -
DR EMBL; AY007375; AAG17947.1; -
DR InterPro; IPR002123; Acyltransferase.
DR Pfam; PF01553; Acyltransferase; 1.
SQ SEQUENCE 583 AA; 66917 MW; 2B2BFAE3E395E049 CRC64;

Query Match 72.7%; Score 40; DB 5; Length 583;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VVPVGMYS 11
: : : : :
Db 227 IIPVGLSVS 235

RESULT 2

Q9BH83 ID Q9BH83 PRELIMINARY; PRT; 583 AA.
AC Q9BH83;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Choline transporter.
 GN SCT1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 ON NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ben Mamoun C., Gluzman I.Y., Goldberg D.E.;
 RT "Plasmodium falciparum choline transporter (PfSCT1) gene."
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY007374; AAK14818.1; -;
 DR EMBL; AY007373; AAK14817.1; -;
 DR InterPro: IPR002123; Acyltransferase.
 DR Pfam: PF01553; Acyltransferase; 1.
 SQ SEQUENCE 583 AA; 66918 MW; 4DF83D7530527474 CRC64;

Query Match 72.7%; Score 40; DB 5; Length 583;
 Best Local Similarity 66.7%; Pred. No. 14;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VVPVGMYS 11
 Db 227 IIPVGLSYS 235

RESULT 3

Q971S2 PRELIMINARY; PRT; 219 AA.
 AC Q971S2
 DT 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Putative ribose 5-phosphate isomerase.
 GN ST1302.
 OS Sulfolobus tokodaii.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 ON NCBI_TaxID=111955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA KAWARABAYASI Y., HINO Y., HORIKAWA H., JIN-NO K., TAKAHASHI M.,
 RA SEKINE M., BABA S.-I., ANKAI A., KOSUGI H., HOSOIYAMA A., FUKUI S.,
 RA NAGAI Y., NISHIJIMA K., OTSUOKA R., NAKAZAWA H., TAKAMIYA M., KATO Y.,
 RA YOSHIZAWA T., TANAKA T., KUDOH Y., YAMAZAKI J., KUSHIDA N., OGUCHI A.,
 RA AOKI K.-I., MASUDA S., YANAGII M., NISHIMURA M., YAMAGISHI A.,
 RA OSHIMA T., KIKUCHI H.;
 RT "Complete genome sequence of an aerobic thermoacidophilic
 Crenarchaeon, Sulfolobus tokodaii strain 7."
 RL DNA Res. 8:123-140(2001).
 DR EMBL; AP000985; BAB66348.1; -;
 DR InterPro: IPR004788; RpiA.
 DR Pfam; PF005813; RpiA; 1.
 SQ SEQUENCE 219 AA; 24541 MW; A4E9A3F2C4006D90 CRC64;

Query Match 70.9%; Score 39; DB 17; Length 219;
 Best Local Similarity 77.8%; Pred. No. 7.4;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 EVVPVGMYS 10
 Db 131 EVVPVGVAY 139

RESULT 4

Q98BP5 PRELIMINARY; PRT; 541 AA.
 AC Q98BP5
 DT 01-OCT-2001 (TRENBLrel. 18, Created)
 DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)

DE Probable DNA ligase.
 GN MLI5481.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.
 ON NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NAFF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kayokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Kouchizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti."
 RL DNA Res. 7:331-338(2000).
 DR EMBL; AP003006; BAB51927.1; -;
 DR InterPro: IPR000977; DNA_ligase.
 DR Pfam; PF01668; DNA_ligase; 1.
 DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN 1.
 DR PROSITE; PS0160; DNA_LIGASE_A3; 1.
 KW Ligase; Complete proteome.
 SQ SEQUENCE 541 AA; 60645 MW; 2EFEF705453F28F8 CRC64;

Query Match 70.9%; Score 39; DB 16; Length 541;
 Best Local Similarity 70.0%; Pred. No. 20;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPVGMYS 10
 Db 445 EELVPVGRAY 454

RESULT 5

Q8RE56 PRELIMINARY; PRT; 209 AA.
 ID Q8RE56
 AC Q8RE56
 DT 01-JUN-2002 (TRENBLrel. 21, Created)
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Hypothetical lipoprotein FN1269.
 GN FN1269.
 OS Fusobacterium nucleatum (subsp. nucleatum).
 OC Bacteria; Fusobacteriia; Fusobacterium.
 ON NCBI_TaxID=76856;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 25586;
 RX MEDLINE=21886394; PubMed=11889109;
 RA Kapral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
 RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
 RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
 RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
 RA Fonstein M., Kyripides N., Overbeek R.;
 RT "Genome sequence and analysis of the oral bacterium Fusobacterium
 nucleatum strain ATCC 25586."
 RL J. Bacteriol. 184:2005-2018(2002).
 DR EMBL; AF010632; AAL95465.1; -;
 KW Lipoprotein; Complete proteome; Hypothetical protein.
 SQ SEQUENCE 209 AA; 24056 MW; E471F6C4911506DA CRC64;

Query Match 69.1%; Score 38; DB 16; Length 209;
 Best Local Similarity 77.8%; Pred. No. 11;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 VVPVGMYS 11
 Db 144 LVPVGLSYS 152

RESULT 6

```

Q97DE7 ID Q97DE7 PRELIMINARY; PRT; 84 AA.
AC Q97DE7;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE Hypothetical protein CAC3530.
GN CAC3530.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Bleton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE007849; AK81456.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 84 AA; 9313 MW; A8E15B62F24DA00B CRC64;

Query Match 67.3%; Score 37; DB 16; Length 84;
Best Local Similarity 70.0%; Pred. No. 6.5;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEVVPVGMYS 10
Db 68 EEVINGVSVY 77

RESULT 7
Q8R8K6 ID Q8R8K6 PRELIMINARY; PRT; 584 AA.
AC Q8R8K6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Predicted ATPases of the HSP70 class involved in cell division.
GN FTS33 OR TTE1990.
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Thermoanaerobacteriales; Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4T / JCM11007;
RX MEDLINE=2192816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
DR EMBL; AE013150; AM25168.1; -.
KW Cell division; Complete proteome.
SQ SEQUENCE 584 AA; 64580 MW; 652BBCC040578F45 CRC64;

Query Match 67.3%; Score 37; DB 16; Length 584;
Best Local Similarity 45.5%; Pred. No. 57;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EEVVPVGMYS 11
Db 368 ESITPIGIAYS 378

RESULT 8
Q8SS39 ID Q8SS39 PRELIMINARY; PRT; 933 AA.
AC Q8SS39;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative ATP-dependent RNA helicase (SKI2 subfamily).
GN ECU04_0910.
OS Encephalitozoon cuniculi.
OC Eukaryota; Microsporidia; Unikaryoniidae; Encephalitozoon.
OX NCBI_TaxID=6035;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RX Genoscope;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RX MEDLINE=21576510; PubMed=11719806;
RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,
RA Prentier G., Barbe V., Peyretailade E., Brottier P., Wincker P.,
RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA Weissenbach J., Vivares C.P.;
RT "Genome sequence and gene compaction of the eukaryote parasite
RT Encephalitozoon cuniculi.";
RL Nature 414:450-453(2001).
DR EMBL; AL590444; CAD25278.1; -.
KW Hypothetical protein.
SQ SEQUENCE 933 AA; 106883 MW; CA400EE3A7AFF7CF CRC64;

Query Match 67.3%; Score 37; DB 5; Length 933;
Best Local Similarity 63.6%; Pred. No. 97;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EEVVPVGMYS 11
Db 43 EAVVPVGAAYT 53

RESULT 9
Q9V7C7 ID Q9V7C7 PRELIMINARY; PRT; 1305 AA.
AC Q9V7C7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CG12961 protein.
GN CG12961.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

```

```

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reiner K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
DR EMBL; AE003810; AAP58131.1; -;
DR FlyBase; FBgn0034024; CGI2961.
SQ SEQUENCE 1305 AA; 152870 MW; 8640B93F47B6EF3F CRC64;

Query Match 67.3%; Score 37; DB 5; Length 1305;
Best Local Similarity 54.5%; Pred. No. 1.4e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVVPGVMSYS 11
|:|:|:|:|:|
Db 1010 EDLVDIGMSYA 1020

RESULT 10
Q12479
ID Q12479 PRELIMINARY; PRT; 156 AA.
AC Q12479;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE ORF YOR013W.
GN YOR013W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA De haan M., Grivell L.A., Maarse A.C.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MIPS;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA STRAIN=FY1679;
RN [4]
RP De haan M., Maarse A.C., Grivell L.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA STRAIN=FY1679;
RX MEDLINE=94019318; PubMed=9413243.
RA Dumont M.E., Schlichter J.B., Cardillo T.S., Hayes M.K., Bethlendy G.,
RA Sherman F.;
RT "CYC2 encodes a factor involved in mitochondrial import of yeast
RT cytochrome c.";
RL Mol. Cell. Biol. 13:6442-6451(1993).
RN [5]
RP SEQUENCE FROM N.A.
RA STRAIN=FY1679;

```

```

RX MEDLINE=94169519; PubMed=7764548;
RA Lee Y.S., Shimizu J., Yoda K., Yamasaki M.;
RT "Molecular cloning of a gene, DHS1, which complements a drug-
RT hypersensitive mutation of the yeast Saccharomyces cerevisiae.";
RL Biosci. Biotechnol. Biochem. 58:391-395(1994).
DR EMBL; Z74920; CAA99201.1; -;
DR EMBL; X87331; CAA60762.1; -;
DR SGD; S000533; YOR013W.
SQ SEQUENCE 156 AA; 17881 MW; 380442B74C272B41 CRC64;

Query Match 65.5%; Score 36; DB 3; Length 156;
Best Local Similarity 66.7%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EVVPGVMSYS 10
|:|:|:|:|
Db 50 EVVPLGWDY 58

RESULT 11
O32330
ID O32330 PRELIMINARY; PRT; 234 AA.
AC O32330;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Sn-glycerol-3-phosphate acyltransferase.
GN PLSD.
OS Clostridium butyricum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1492;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98053837; PubMed=9393688;
RA Heath R.J., Goldfine H., Rock C.O.;
RT "A gene (plsB) from Clostridium butyricum that functionally
RT substitutes for the sn-glycerol-3-phosphate acyltransferase gene
RT (plsB) of Escherichia coli.";
RL J. Bacteriol. 179:7257-7263(1997).
DR EMBL; AF009362; AAC46006.1; -;
DR InterPro; IPR002123; Acyltransferase.
DR Pfam; PF01553; Acyltransferase; 1.
KW Acyltransferase; Transferase.
SQ SEQUENCE 234 AA; 26198 MW; A4C15861A2995C37 CRC64;

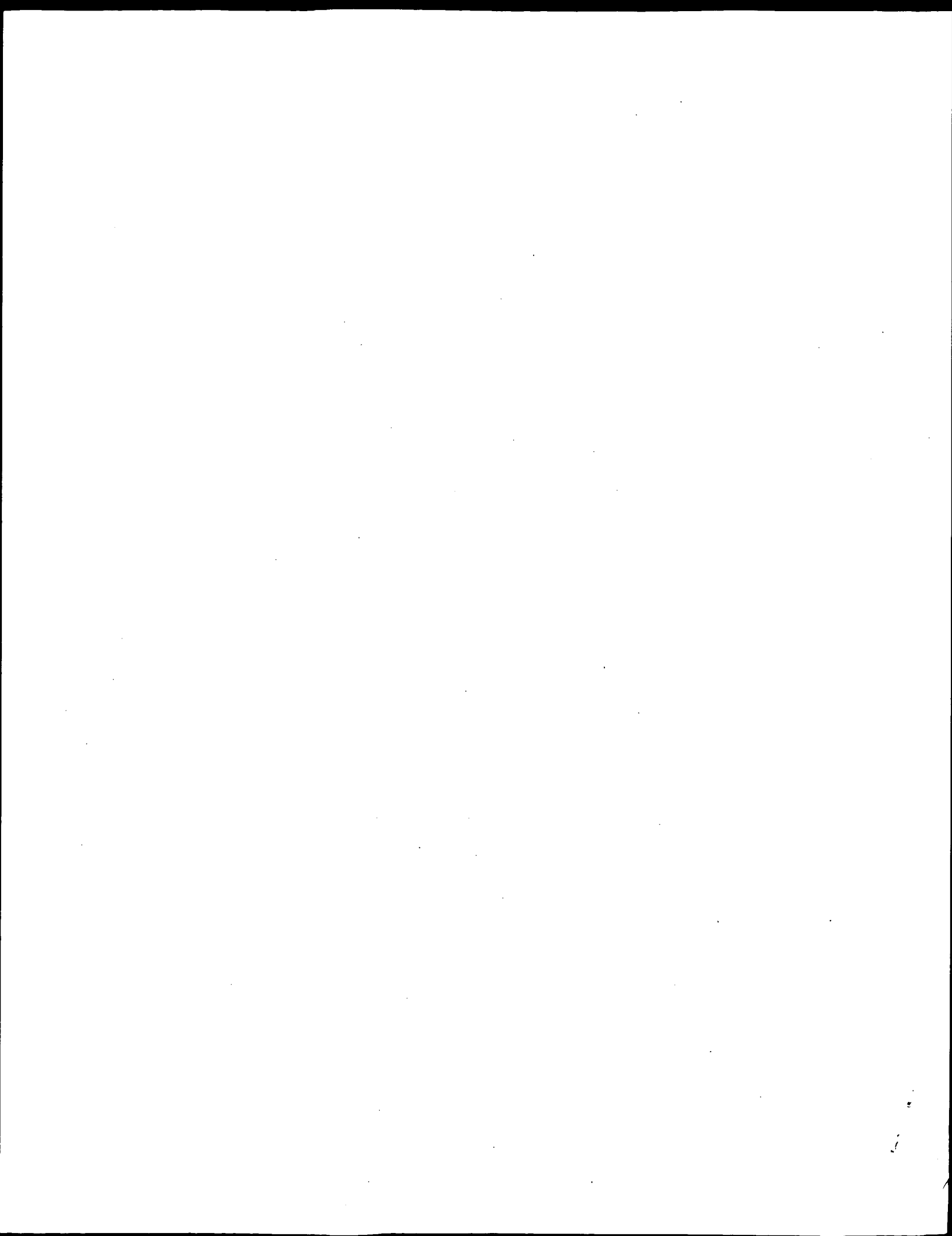
Query Match 65.5%; Score 36; DB 2; Length 234;
Best Local Similarity 62.5%; Pred. No. 33;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EVVPGVMS 9
|:|:|:|:|
Db 156 EIIPIGMS 163

RESULT 12
O9VWV3
ID O9VWV3 PRELIMINARY; PRT; 273 AA.
AC O9VWV3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE MTPB protein.
GN MTPB.
OS Mycobacterium avium.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1764;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=2151;
RA Eckstein T.M., Lambert M.L., Brennan P.J., Belisle J.T., Inamine J.M.;
RT "Identification of a gene cluster involved in glycopeptidolipid

```

RT biosynthesis and of a gene cluster encoding daunorubicin resistance in two strains of *Mycobacterium avium* serovar 2.";
Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:08:49 ; Search time 11 Seconds
(without alignments)
96.134 Million cell updates/sec

Title: AUDET-909-4

Perfect score: 55

Sequence: 1 eevvpvmsys 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_73.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	69.1	1498	2 B97355	DNA segregation AT
2	37	67.3	84	2 E97333	hypothetical prote
3	36	65.5	156	2 S54619	hypothetical prote
4	36	65.5	433	1 S48515	adenylosuccinate s
5	36	65.5	488	2 A82984	hypothetical prote
6	36	65.5	541	2 AH2679	ATP-dependent DNA
7	36	65.5	573	2 F97461	DNA ligase (AB0425
8	36	65.5	664	2 A72273	hypothetical prote
9	36	65.5	827	2 A41672	glycerol-3-phospha
10	36	65.5	840	2 T93116	probable sulfate p
11	36	65.5	877	2 T40413	sulfate permease -
12	36	65.5	3472	2 T31308	hypothetical 367K
13	35	63.6	116	2 E90544	50S ribosomal prot
14	35	63.6	227	2 E75619	hypothetical prote
15	35	63.6	286	1 C42053	gap junction prote
16	35	63.6	670	2 S22293	zinc finger protei
17	35	63.6	919	2 T16459	hypothetical prote
18	35	63.6	2717	2 A34203	DNA-binding protei
19	34	61.8	102	2 A42452	V1 protein - tobac
20	34	61.8	252	2 H69491	cell division inhi
21	34	61.8	271	2 A36893	transcription acti
22	34	61.8	288	2 JC4011	cyclin D2 - rat
23	34	61.8	288	2 I58372	cyclin D2 - mouse
24	34	61.8	289	2 A41984	cyclin D2 - human
25	34	61.8	289	2 A42822	cyclin D1 - Africa
26	34	61.8	291	2 S57922	cyclin D2 - Africa
27	34	61.8	291	2 S57925	cyclin D2 - chicke
28	34	61.8	291	2 JC4579	cyclin D2 - zebra
29	34	61.8	291	2 S62730	

30 34 61.8 292 2 B42822 cyclin D3 - human
31 34 61.8 295 2 A38977 cyclin D1 - human
32 34 61.8 295 2 A56523 cyclin D1 - mouse
33 34 61.8 295 2 JC2342 cyclin D1 - rat
34 34 61.8 317 2 C85432 hypothetical prote
35 34 61.8 327 2 D71278 hypothetical prote
36 34 61.8 521 2 T41621 hypothetical prote
37 34 61.8 529 2 D72253 hypothetical prote
38 34 61.8 692 1 S46953 phosphotransferase
39 34 61.8 944 2 S01909 hairy wing suppres
40 34 61.8 1150 2 T20173 hypothetical prote
41 33 60.0 225 2 S57810 hypothetical prote
42 33 60.0 227 2 F90249 ribose 5-phosphate
43 33 60.0 240 2 AB0338 probable membrane
44 33 60.0 258 1 A64451 probable 3-isoprop
45 33 60.0 259 2 T34536 hypothetical prote

ALIGNMENTS

RESULT 1

B97355

DNA segregation ATPase, FtsK/SpoIIIE family, YUKA B. subtilis ortholog [imported] - Clost
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C;Accession: B97355
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: B97355
A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1498 <KUR>

A;Cross-references: GB:AE001437; PIDN:AK81629.1; PID:gl5026814; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC3709

Query Match 69.1%; Score 38; DB 2; Length 1498;
Best Local Similarity 60.0%; Pred. No. 54;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPVGMYSY 10

Db 1276 EQKIPGMYSY 1285

RESULT 2

E97333

hypothetical protein CAC3530 [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C;Accession: E97333
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: E97333
A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-84 <KUR>

A;Cross-references: GB:AE001437; PIDN:AK81456.1; PID:gl5026624; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC3530

Query Match 67.3%; Score 37; DB 2; Length 84;

Best Local Similarity 70.0%; Pred. No. 3.8;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVVPVGMYS 10
 |||: |||:
 Db 68 EEVINGVSV 77

RESULT 3

S54619
 hypothetical protein YOR013w - yeast (Saccharomyces cerevisiae)
 N;Alternate names: hypothetical protein O2612; hypothetical protein VOL303.3
 C;Species: Saccharomyces cerevisiae
 C;Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 19-Apr-2002
 C;Accession: S54619; S66879
 R;de Haan, M.; Maarse, A.C.; Grivell, L.A.
 submitted to the EMBL Data Library, May 1995
 A;Reference number: S54617
 A;Accession: S54619
 A;Molecule type: DNA
 A;Residues: 1-156 <DEH>
 A;Cross-references: EMBL:X87331; NID:g1041652; PIDN:CAA60762.1; PID:g829123
 R;de Haan, M.; Grivell, L.A.; Maarse, A.C.
 submitted to the Protein Sequence Database, July 1996
 A;Reference number: S66877
 A;Accession: S66879
 A;Molecule type: DNA
 A;Residues: 1-156 <DEW>
 A;Cross-references: EMBL:Z74920; NID:g1420109; PIDN:CAA99201.1; PID:g1420111; MIPS:YOR013w
 A;Experimental source: strain S288C
 C;Genetics:
 A;Cross-references: SGD:S0005539
 A;Map position: 15R
 C;Superfamily: hypothetical protein YOR013w

Query Match 65.5%; Score 36; DB 2; Length 156;
 Best Local Similarity 66.7%; Pred. No. 12;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPVGMYS 10
 |||: |||:
 Db 50 EVVPLGMDY 58

RESULT 4

S48515
 adenylosuccinate synthase (EC 6.3.4.4) - yeast (Saccharomyces cerevisiae)
 N;Alternate names: protein N1290; protein YNL220w
 C;Species: Saccharomyces cerevisiae
 C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
 C;Accession: S48515; S53085; S63178
 R;Shabes, A.V.; Andreichuk, Y.V.; Holmes, W.M.; Domkin, V.D.
 submitted to the EMBL Data Library, July 1993
 A;Reference number: S48515
 A;Accession: S48515
 A;Molecule type: DNA
 A;Residues: 1-433 <SHA>
 A;Cross-references: EMBL:L22185; NID:g347862; PIDN:AAA91338.1; PID:g347863
 R;Ohanjan, T.; Daigman-Fornier, B.; Krauss, G.
 submitted to the EMBL Data Library, March 1995
 A;Description: Adenylosuccinate synthetase from Saccharomyces cerevisiae.
 A;Reference number: S53085
 A;Accession: S53085
 A;Molecule type: DNA
 A;Residues: 1-433 <SHA>
 A;Cross-references: EMBL:Z48671; NID:g732938; PIDN:CAA88590.1; PID:g732939
 R;Dueserboett, A.; Floeth, M.; Fritz, C.; Heuss-Neitzel, D.; Hilbert, H.; Moestl, D.
 submitted to the Protein Sequence Database, April 1996
 A;Reference number: S62944
 A;Accession: S63178
 A;Molecule type: DNA
 A;Residues: 1-433 <DUE>
 A;Cross-references: EMBL:Z71496; NID:g1302236; PIDN:CAA96123.1; PID:g1302237; GSPDB:GN00186
 A;Experimental source: strain S288C
 C;Genetics:

A;Gene: SGD:ADE12; MIPS:YNL220w
 A;Cross-references: SGD:S0005164; MIPS:YNL220w
 A;Map position: 14L
 C;Superfamily: adenylosuccinate synthase
 C;Keywords: ligase; purine nucleotide biosynthesis

Query Match 65.5%; Score 36; DB 1; Length 433;
 Best Local Similarity 75.0%; Pred. No. 35;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 VPVGMYS 11
 :|||: |||:
 Db 347 IPVGISYS 354

RESULT 5

A82984
 hypothetical protein PA5294 [imported] - Pseudomonas aeruginosa (strain PA01)
 C;Species: Pseudomonas aeruginosa
 C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 24-Aug-2001
 C;Accession: A82984
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Brj
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
 A;Reference number: A82950; MUID:20437337; PMID:10984043
 A;Accession: A82984
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-488 <STO>
 A;Cross-references: GB:AE004942; GB:AE004091; NID:g9951607; PIDN:AAG08679.1; GSPDB:GN001
 C;Genetics:
 A;Experimental source: strain PA01
 A;Gene: PA5294
 C;Superfamily: conserved hypothetical protein H11612

Query Match 65.5%; Score 36; DB 2; Length 488;
 Best Local Similarity 66.7%; Pred. No. 40;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VPVGMYS 11
 :|||: |||:
 Db 286 MVPVGLSYA 294

RESULT 6

AH2679
 ATP-dependent DNA ligase Atu0840 [imported] - Agrobacterium tumefaciens (strain C58, Dup
 C;Species: Agrobacterium tumefaciens
 C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
 C;Accession: AH2679
 R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClella
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E
 ster, E.W.
 A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A;Reference number: AB2577; PMID:11743193
 A;Accession: AH2679
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-541 <KUR>
 A;Cross-references: GB:AE008688; PIDN:AAL41854.1; PID:g17739214; GSPDB:GN00186
 A;Experimental source: strain C58 (Dupont)
 C;Genetics:
 A;Gene: Atu0840
 A;Map position: circular chromosome

Query Match 65.5%; Score 36; DB 2; Length 541;
 Best Local Similarity 60.0%; Pred. No. 45;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPVGMYSY 10
 : : : : :
 Db 442 EQLVPVGKAY 451

RESULT 7
 F97461
 DNA ligase (AB042527) [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
 C:Species: Agrobacterium tumefaciens
 C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
 C:Accession: F97461
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurrello, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
 A:Reference number: A97359; PMID:11743194
 A:Accession: F97461
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-573 <KUR>
 A:Cross-references: GB:AE007869; PIDN:AAK86647.1; PID:g15155825; GSPDB:GN00169
 C:Genetics:
 A:Gene: AGR_C_1536
 A:Map position: circular chromosome

Query Match 65.5%; Score 36; DB 2; Length 573;
 Best Local Similarity 60.0%; Pred. No. 48;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPVGMYSY 10
 : : : : :
 Db 474 EQLVPVGKAY 483

RESULT 8
 A72273
 hypothetical protein TM1277 - Thermotoga maritima (strain MSB8)
 C:Species: Thermotoga maritima
 C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
 C:Accession: A72273
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
 C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
 A:Reference number: A72200; PMID:99287316; PMID:10360571
 A:Accession: A72273
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-664 <ARN>
 A:Cross-references: GB:AE001783; GB:AE000512; NID:g4981832; PIDN:AAD36352.1; PID:g498183
 A:Experimental source: strain MSB8
 C:Genetics:
 A:Gene: TM1277

Query Match 65.5%; Score 36; DB 2; Length 664;
 Best Local Similarity 54.5%; Pred. No. 56;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPVGMYSY 11
 : : : : :
 Db 359 EYITPVGIAYS 369

RESULT 9
 A41672
 glycerol-3-phosphate acyltransferase homolog - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 05-Nov-1999
 C:Accession: A41672
 R:Shin, D.H.; Paulauskis, J.D.; Moustaid, N.; Sul, H.S.
 J. Biol. Chem. 266, 23834-23839, 1991

A:Title: Transcriptional regulation of p90 with sequence homology to Escherichia coli g1
 A:Reference number: A41672; PMID:92084678; PMID:1721057
 A:Accession: A41672
 A:Molecule type: mRNA
 A:Residues: 1-827 <SHI>
 A:Cross-references: GB:M77003; NID:g193366; PIDN:AAA37647.1; PID:g193367
 A:Experimental source: liver

Query Match 65.5%; Score 36; DB 2; Length 827;
 Best Local Similarity 75.0%; Pred. No. 72;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 VVPVGMYSY 10
 : : : : :
 Db 349 VIPVGISY 356

RESULT 10

T39116

probable sulfate permease - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T39116

R:Hunt, C.; Aves, S.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, November 1999

A:Reference number: Z21829

A:Accession: T39116

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-840 <HUN>

A:Cross-references: EMBL:AL132779; PIDN:CA860015.1; GSPDB:GN00066; SPDB:SPAC869.05C

A:Experimental source: strain 972h; cosmid c869

C:Genetics:

A:Gene: SPDB:SPAC869.05C

A:Map position: 1

Query Match 65.5%; Score 36; DB 2; Length 840;
 Best Local Similarity 77.8%; Pred. No. 73;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPVGMYSY 11
 : : : : :
 Db 135 VVPQGSYA 143

RESULT 11

T40413

sulfate permease - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T40413

R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Jimenez Martinez, J.

submitted to the EMBL Data Library, August 1998

A:Reference number: Z21926

A:Accession: T40413

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-877 <LYN>

A:Cross-references: EMBL:AL031261; PIDN:CAA20298.1; GSPDB:GN00067; SPDB:SPBC3H7.02

A:Experimental source: strain 972h; cosmid C3H7

C:Genetics:

A:Gene: SPDB:SPBC3H7.02

A:Map position: 2

Query Match 65.5%; Score 36; DB 2; Length 877;
 Best Local Similarity 77.8%; Pred. No. 76;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPVGMYSY 11
 : : : : :
 Db 148 VVPQGSYA 156

RESULT 12

T31308
hypothetical 367K protein - Cenarchaeum symbiosum
C:Species: Cenarchaeum symbiosum
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 18-Feb-2000
C:Accession: T31308
R:Schleper, C.; DeLong, E.F.; Preston, C.M.; Feldman, R.A.; Wu, K.Y.; Swanson, R.V.
J. Bacteriol. 180, 5003-5009, 1998
A:Title: Genomic analysis reveals chromosomal variation in natural populations of the un
A:Reference number: Z20994; MUID:98422450; PMID:9748430
A:Accession: T31308
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-3472 <SCH>
A:Cross-references: EMBL:AF083072; NID:G3599393; PID:G3599394; PIDN:AAAC62699.1
C:Superfamily: Cenarchaeum symbiosum hypothetical 367K protein

```
Query Match          65.5%; Score 36; DB 2; Length 3472;
Best Local Similarity 54.5%; Pred. No. 3.4e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      1 EEVVPVGM SYS 11
         | : | : | : | : |
Db     2294 EDVIPRGISFS 2304
```

RESULT 13

E90544 13
50S ribosomal protein L20 [imported] - Mycoplasma pulmonis (strain UAB CT1P)
C:Species: Mycoplasma pulmonis
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C:Accession: E90544
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
Nucleic Acids Res. 29, 245-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmonis
A:Reference number: A95912; MUID:21267165; PMID:11353084

Genetic Code: 5953
 Superfamily: Escherichia coli ribosomal protein L20

Query Match	63.6%	Score 35;	DB 2;	Length 116;
Best Local Similarity	77.8%;	Pred. No. 13;		
Matches	7;	Conservative	1;	Mismatches
			1;	Indels
			0;	Gaps
			0;	

Qy	3	VPV	GMSYS	11
			:	
Db	68	VRPL	GMSYS	76

RESULT 14

hypothetical protein DRB0013 - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: E75619
R:Whice, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

A; Experimental source: strain R1
C; Genetics:
A; Gene: DRB0013
A; Map position: megaplasmid
A; Genome: plasmid
A; Note: plasmid MP1
C; Superfamily: *Peinococcus radiodurans* megaplasmid
A; Superfamily: hypothetical protein DRB0013

Query Match	63.6%;	Score 35;	DB 2;	Length 227;
Best Local Similarity	54.5%;	Pred. NO. 28;		
Matches 6;	Conservative	3;	Mismatches	2;
			Indels	0;
			Gaps	0;

QY 1 EEVVPVGMSSYS 11
Db 43 ESVLPIGHSSFS 53

RESULT 15

gap junction protein Cx33 - rat
N; Alternate names: connexin 33
C; Species: Rattus norvegicus (Norway rat)
C; Date: 04-Mar-1993 #sequence_revision 02-Jun-1994 #text_change 22-Jun-1999
C; Accession: C42053
R; Haefliger, J.A.; Bruzone, R.; Jenkins, N.A.; Gilbert, D.J.; Copeland, N.G.; Paul, D.L.
J. Biol. Chem. 267, 2057-2064, 1992
A; Title: Four novel members of the connexin family of gap junction proteins. Molecular c
Reference number: A42053; MUID:92112940; PMID:1370487

A;Accession: C12033
A;Molecule type: DNA
A;Residues: 1-286 <HAE>
A;Cross-references: GB:M76534; NID:G203663; PID:AAA0998.1; PIR:G12033
A;Note: sequence extracted from NCBI backbone (NCBIP:76095)
C;Superfamily: gap junction protein
C;Keywords: gap junction; phosphoprotein; transmembrane protein
F;1-23/Domain: intracellular #status predicted <INT1>
F;24-41/Domain: transmembrane #status predicted <TM1>
F;42-77/Domain: extracellular #status predicted <E1>
F;78-97/Domain: transmembrane #status predicted <TM2>
F;98-150/Domain: intracellular #status predicted <INT2>
F;151-187/Domain: transmembrane #status predicted <TM3>
F;188-207/Domain: extracellular #status predicted <E2>
F;208-234/Domain: transmembrane #status predicted <TM4>
F;235-286/Domain: intracellular #status predicted <INT3>

Query Match	63.6%;	Score 35;	DB 1;	Length 286;
Best Local Similarity	66.7%;	Pred. No. 36;		
Matches	6;	Conservative	3;	Mismatches
			0;	Indels
			0;	Gaps

QY 1 EEVVPVGM 9
: : : : : :
db 274 DOWPVGLS 282

Search completed: June 4, 2003, 13:14:16
Job time : 11 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:04:34 ; Search time 6.25 Seconds
(without alignments)
72.998 Million cell updates/sec

Title: AUDET-909-4

Perfect score: 55

Sequence: 1 eevrvpmsys 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	38	69.1	1498	1	YIA9_CLOAB
2	36	65.5	432	1	PURA_YEAST
3	36	65.5	488	1	NOM2_PSEAE
4	36	65.5	827	1	PLSB_MOUSE
5	36	65.5	828	1	PLSB_RAT
6	36	65.5	877	1	SULH_SCHPO
7	36	65.5	1058	1	CARB_FUSNN
8	35	63.6	116	1	RL20_MYCPU
9	35	63.6	286	1	CXAB_RAT
10	35	63.6	828	1	PLSB_HUMAN
11	35	63.6	2717	1	ZEP1_HUMAN
12	34	61.8	102	1	Y11K_TYDVA
13	34	61.8	271	1	POBR_ACICA
14	34	61.8	288	1	CGD2_RAT
15	34	61.8	289	1	CGD2_HUMAN
16	34	61.8	289	1	CGD2_MOUSE
17	34	61.8	291	1	CGD1_BRARE
18	34	61.8	291	1	CGD1_XENLA
19	34	61.8	291	1	CGD2_CHICK
20	34	61.8	291	1	CGD2_XENLA
21	34	61.8	292	1	CGD1_CHICK
22	34	61.8	292	1	CGD3_HUMAN
23	34	61.8	295	1	CGD3_HUMAN
24	34	61.8	295	1	CGD1_MOUSE
25	34	61.8	295	1	CGD1_RAT
26	34	61.8	691	1	QAT6_HUMAN
27	34	61.8	944	1	SUHW_DROME
28	33	60.0	258	1	YC10_METYA
29	33	60.0	349	1	PHOE_KLEBOX
30	33	60.0	350	1	PHOE_ENTCL
31	33	60.0	351	1	PHOE_ECOLI
32	33	60.0	421	1	ACDM_HUMAN
33	33	60.0	421	1	ACDM_FIG

34	33	60.0	426	1	AROA_VIBCH
35	33	60.0	478	1	GSR2_HUMAN
36	33	60.0	521	1	CYOA_HAEIN
37	33	60.0	660	1	CAO1_HUMAN
38	33	60.0	661	1	CAO1_CAVPO
39	33	60.0	661	1	CAO1_MOUSE
40	33	60.0	661	1	CAO1_RAT
41	33	60.0	670	1	QATP_RAT
42	33	60.0	734	1	SYTC_YEAST
43	33	60.0	759	1	SCT1_YEAST
44	33	60.0	762	1	MUS2_AQUAE
45	33	60.0	788	1	CV14_NEUCR

ALIGNMENTS

RESULT 1

YIA9_CLOAB
ID YIA9_CLOAB STANDARD; PRT; 1498 AA.

AC Q04351;

DT 01-FEB-1994 (Rel. 28, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein CAC3709.

GN CAC3709.

OS Clostridium acetobutylicum.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.

OX NCBI_TaxID=1488;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;

RX MEDLINE=21359325; PubMed=11466286;

RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,

Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,

Taturov R.L., Sabathe F., Doucet-Stamm L., Soucaille P., Daly M.J.,

Bennett G.N., Koonin E.V., Smith D.R.;

RT "Genome sequence and comparative analysis of the solvent-producing

Bacterium Clostridium acetobutylicum.";

RL J. Bacteriol. 183:4823-4838(2001).

RN [2]

RP SEQUENCE OF 1-108 FROM N.A.

RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;

RX MEDLINE=93273706; PubMed=8501044;

RA Sauer U., Duerre P.;

RT "Sequence and molecular characterization of a DNA region encoding a

small heat shock protein of Clostridium acetobutylicum.";

RL J. Bacteriol. 175:3394-3400(1993).

CC -!- SIMILARITY: BELONGS TO THE FTSK/SPOIIE FAMILY.

CC -!- CAUTION: Ref.2 sequence differs from that shown due to frameshifts

in positions 76 and 106.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

DR EMBL; AB007866; AAK81629.1; -.

DR EMBL; X65276; CAA46379.1; ALT FRAME.

DR InterPro; IPR002543; FTSK_SpoIIE.

DR Pfam; PF01580; FtsK_SpoIIE; 2.

KW Hypothetical protein, ATP-binding; Complete proteome.

FT NP_BIND 675 682 ATP (POTENTIAL).

SQ SEQUENCE 1498 AA; 168968 MW; FF42037A335A9649 CRC64;

Query Match

Best Local Similarity 69.1%; Score 38; DB 1; Length 1498;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 EVVPMGMSY 10
Db 1276 EQIPMGMSY 1285

Query Match 65.5%; Score 36; DB 1; Length 432;
Best Local Similarity 75.0%; Pred. No. 15;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 VPVGMYSYS 11
Db 346 IPVGISYS 353

RESULT 2
PURA_YEAST STANDARD; PRT; 432 AA.
AC P80210;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Adenylosuccinate synthetase (EC 6.3.4.4) (IMP--aspartate ligase)
DE (ADSS) (AMPSase).
GN ADE12 OR YNL20W OR N1290.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Shabes A.V., Andreichuk Y.V., Holmes W.M., Domkin V.D.;
RL Submitted (JUL-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Ohanjan T., Daignan-Fornier B., Krauss G.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Dueterhoeft A., Floeth M., Fritz C., Heuss-Neitzel D., Hilbert H.,
RL Moestl D.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 1-22 AND 233-244.
RX MEDLINE=9338587; PubMed=8376380;
RA Zeidler R., Hobert O., Johannes L., Paulhammer H., Krauss G.;
RT "Characterization of two novel single-stranded DNA-specific
autonomously replicating sequence-binding proteins from Saccharomyces
cerevisiae, one of which is adenylosuccinate synthetase."
RL J. Biol. Chem. 268:20191-20197(1993).
CC -!- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE DE NOVO PATHWAY OF PURINE
NUCLEOTIDE BIOSYNTHESIS.
CC NUCLEOTIDE ACTIVITY: GTP + IMP + L-aspartate = GDP + phosphate +
adenylosuccinate.
CC -!- PATHWAY: AMP biosynthesis; first committed step.
CC -!- SIMILARITY: BELONGS TO THE ADENYLOSUCCINATE SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; L22185; AA91338.1; -.
DR EMBL; Z48671; CAA98590.1; -.
DR EMBL; Z71496; CAA96123.1; -.
DR HSSP; PL2283; 1ADE.
DR SGD; S0005164; ADE12.
DR InterPro; IPR001114; Asucc_synthetase.
DR Pfam; PF00709; Adenylosucc_synth; 1.
DR ProDom; PD001188; Asucc_synthetase; 1.
DR TIGRFAMs; TIGR00184; purA; 1.
DR PROSITE; PS00513; ADENYLOSUCCIN SYN 2; 1.
DR PROSITE; PS01266; ADENYLOSUCCIN SYN 1; 1.
KW Purine biosynthesis; Ligase; GTP-binding.
FT INIT MET 0
FT NP_BIND 10 16 GTP (BY SIMILARITY).
FT ACT_SITE 144 144 BY SIMILARITY.
FT ACT_SITE 151 151 BY SIMILARITY.
FT CONFLICT 236 236 D -> G (IN REF. 4).
SQ SEQUENCE 432 AA; 48148 MW; FFEB44F46349570 CRC64;

Query Match 65.5%; Score 36; DB 1; Length 488;

QY 4 VPVGMYSYS 11
Db 346 IPVGISYS 353

RESULT 3
NOM2_PSEAE STANDARD; PRT; 488 AA.
AC Q9HTRO;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable multidrug resistance protein norm 2 (Na(+)/drug antiporter)
DE (Multidrug-efflux transporter).
GN NORM2 OR PA5294.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
opportunistic pathogen."
RL Nature 406:959-964(2000).
CC -!- FUNCTION: FUNCTIONS AS A NA(+)/DRUG ANTIPORTER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
(potential).
CC -!- SIMILARITY: BELONGS TO THE MULTI ANTIMICROBIAL EXTRUSION (MATE)
FAMILY. NORM (TC 2.A.66.1.1) SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; AB004942; AAG08679.1; -.
DR InterPro; IPR002528; MateB.
DR Pfam; PF01554; UPF0013; 2.
DR TIGRFAMs; TIGR00797; mateB; 1.
DR Transports; Antiport; Sodium transport; Transmembrane; Inner membrane;
Complete proteome.
FT TRANSMEM 11 31 POTENTIAL.
FT TRANSMEM 55 75 POTENTIAL.
FT TRANSMEM 97 117 POTENTIAL.
FT TRANSMEM 127 147 POTENTIAL.
FT TRANSMEM 159 179 POTENTIAL.
FT TRANSMEM 190 210 POTENTIAL.
FT TRANSMEM 247 267 POTENTIAL.
FT TRANSMEM 271 291 POTENTIAL.
FT TRANSMEM 317 337 POTENTIAL.
FT TRANSMEM 355 375 POTENTIAL.
FT TRANSMEM 393 413 POTENTIAL.
FT TRANSMEM 421 441 POTENTIAL.
FT DOMAIN 82 89 POLY-ALA.
SQ SEQUENCE 488 AA; 50558 MW; 8202CBB392C0F9FD CRC64;

Query Match 65.5%; Score 36; DB 1; Length 488;

```

Best Local Similarity 66.7%; Pred. No. 17;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VVPVQMSYS 11
:||||:|
Db 286 MVPVGLSYA 294

RESULT 4

PLSB MOUSE
ID PLSB_MOUSE STANDARD; PRT; 827 AA.
AC Q61586;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycerol-3-phosphate acyltransferase, mitochondrial precursor
DE (EC 2.3.1.15) (GPAT) (P90).
GN GPAM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92084678; PubMed=1721057;
RA Shin D.-H., Paulauskis J.D., Moustaid N., Sul H.S.;
RT "Transcriptional regulation of p90 with sequence homology to
RT Escherichia coli glycerol-3-phosphate acyltransferase.";
RL J. Biol. Chem. 266:23834-23839(1991).
CC -!- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
CC acyl-sn-glycerol 3-phosphate.
CC -!- PATHWAY: FIRST STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS. IT MAY
CC ALSO FUNCTION IN THE REGULATION OF MEMBRANE BIOGENESIS.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC outer membrane (by similarity).
CC -!- TISSUE SPECIFICITY: HIGHEST LEVELS IN LIVER, INTERMEDIATE LEVELS
CC IN MUSCLE AND KIDNEY, AND LOWEST LEVELS IN LUNG AND BRAIN.
CC -!- SIMILARITY: BELONGS TO THE GPAT / DAPAT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M77003; AAA37647.1; -.
CC DR MGD; MG1:109162; Gpam.
CC DR InterPro; IPR002123; Acyltransferase.
CC DR Pfam; PF01553; Acyltransferase; 1.
CC KW Phospholipid biosynthesis; Transferase; Acyltransferase;
CC Transmembrane; Mitochondrion; Transit peptide.
CC TRANSIT 1 ?
CC CHAIN ? 827 MITOCHONDRION (POTENTIAL).
CC FT DOMAIN ? 471 MITOCHONDRIAL INTERMEMBRANE (POTENTIAL).
CC FT DOMAIN 472 494 POTENTIAL.
CC FT TRANSMEM 495 574 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 575 593 POTENTIAL.
CC FT TRANSMEM 594 827 MITOCHONDRIAL INTERMEMBRANE (POTENTIAL).
CC FT DOMAIN 827 AA; 93689 MW; A64EDB697BD664B3 CRC64;
CC SEQUENCE

Query Match 65.5%; Score 36; DB 1; Length 827;
Best Local Similarity 75.0%; Pred. No. 29;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 VVPVQMSYS 10
:||||:|
Db 349 VIPVGISY 356

RESULT 5

Best Local Similarity 75.0%; Pred. No. 29;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

PLSB RAT
ID PLSB_RAT STANDARD; PRT; 828 AA.
AC P97564; P97565; O35349;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Glycerol-3-phosphate acyltransferase, mitochondrial precursor
DE (EC 2.3.1.15) (GPAT).
GN GPAM.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND ACTIVITY.
RC TISSUE=Liver;
RX MEDLINE=99376617; PubMed=10446428;
RA Bhat B.G., Wang P., Kim J.-H., Black T.M., Lewin T.M.,
RA Fiedorek F.T. Jr., Coleman R.A.;
RT "Rat sn-glycerol-3-phosphate acyltransferase: molecular cloning and
RT characterization of the cDNA and expressed protein.";
RL Biochim. Biophys. Acta 1439:415-423(1999).
CC [2]
CC SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TOPOLOGY.
CC STRAIN=Sprague-Dawley; TISSUE=Liver;
CC RX MEDLINE=20493538; PubMed=10924502;
CC RA Ballija V.S., Chakraborty T.R., Nikonov A.V., Morimoto T., Haldar D.;
CC RT "Identification of two transmembrane regions and a cytosolic domain of
CC rat mitochondrial glycerophosphate acyltransferase.";
CC RL J. Biol. Chem. 275:31668-31673(2000).
CC -!- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
CC acyl-sn-glycerol 3-phosphate.
CC -!- PATHWAY: FIRST STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS. IT MAY
CC ALSO FUNCTION IN THE REGULATION OF MEMBRANE BIOGENESIS.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC outer membrane.
CC -!- SIMILARITY: BELONGS TO THE GPAT / DAPAT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF021348; AAB71605.1; -.
CC DR EMBL; U36771; AAB39470.2; ALT_INIT.
CC DR InterPro; IPR002123; Acyltransferase.
CC DR Pfam; PF01553; Acyltransferase; 1.
CC KW Phospholipid biosynthesis; Transferase; Acyltransferase;
CC Transmembrane; Mitochondrion; Transit peptide.
CC TRANSIT 1 ?
CC CHAIN ? 828 MITOCHONDRION (POTENTIAL).
CC FT DOMAIN ? 471 MITOCHONDRIAL INTERMEMBRANE (POTENTIAL).
CC FT TRANSMEM 472 494 POTENTIAL.
CC FT DOMAIN 495 574 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 575 593 POTENTIAL.
CC FT DOMAIN 594 828 MITOCHONDRIAL INTERMEMBRANE (POTENTIAL).
CC FT TRANSIT 37 37
CC FT CONFLICT 35 85
CC FT CONFLICT 300 300 C -> F (IN REF. 2).
CC FT CONFLICT 331 331 I -> V (IN REF. 2).
CC FT CONFLICT 331 331 L -> V (IN REF. 2).
CC FT CONFLICT 472 474 ILF -> NL (IN REF. 2).
CC FT CONFLICT 497 497 R -> W (IN REF. 2).
CC FT CONFLICT 602 606 SAGGL -> LPEP (IN REF. 2).
CC FT CONFLICT 644 644 Q -> H (IN REF. 2).
CC FT CONFLICT 744 744 G -> A (IN REF. 2).
CC FT CONFLICT 744 744
CC SEQUENCE 828 AA; 93714 MW; ACA4A087E7DEB12C CRC64;
CC

Query Match 65.5%; Score 36; DB 1; Length 828;
Best Local Similarity 75.0%; Pred. No. 29;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 VVPVMSY 10
|:|:|:|:
Db 349 VIPVGSY 356

RESULT 6
SULH SCHPO STANDARD; PRT; 877 AA.
AC 074377;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable sulfate permease C3H7.02.
GN SPBC3H7.02.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hanlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leherach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Lambert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
RT "The genome sequence of Schizosaccharomyces pombe";
RL Nature 415:871-880(2002).
CC -1- FUNCTION: HIGH AFFINITY UPTAKE OF SULFATE INTO THE CELL (BY
SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE SLC26A FAMILY OF TRANSPORTERS.
CC -1- SIMILARITY: CONTAINS 1 STAS DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL031261; CAA20298.1;
DR InterPro; IPR002645; STAS
DR InterPro; IPR001902; Sulfate_transp.
DR Pfam; PF00916; Sulfate_transp; 1.
DR Pfam; PF01740; STAS; 1.
DR TIGRFAMs; TIGR00815; sulP; 1.
DR PROSITE; PS01130; SLC26A; 1.
DR PROSITE; PS0801; STAS; 1.
KW Transport; Transmembrane.

FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 161 181 POTENTIAL.
FT TRANSMEM 186 206 POTENTIAL.
FT TRANSMEM 221 241 POTENTIAL.
FT TRANSMEM 243 263 POTENTIAL.
FT TRANSMEM 292 312 POTENTIAL.
FT TRANSMEM 329 349 POTENTIAL.
FT TRANSMEM 384 404 POTENTIAL.
FT TRANSMEM 424 444 POTENTIAL.
FT TRANSMEM 461 481 POTENTIAL.
FT TRANSMEM 484 504 POTENTIAL.
FT TRANSMEM 518 538 POTENTIAL.
FT TRANSMEM 543 563 POTENTIAL.
FT DOMAIN 594 747 STAS.
SQ SEQUENCE 877 AA; 96373 MW; 56995A8493371B43 CRC64;
Query Match 65.5%; Score 36; DB 1; Length 877;
Best Local Similarity 77.8%; Pred. No. 31;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 3 VVPVMSYS 11
|:|:|:|:
Db 148 VVPQMSYA 156

RESULT 7
CARB_FUSNN STANDARD; PRT; 1058 AA.
AC Q8RG86;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-
phosphate synthetase ammonia chain).
DE CARB OR FN0422.
GN Fusobacterium nucleatum (subsp. nucleatum).
OS Fusobacterium nucleatum; Fusobacterium.
OC Bacteria; Fusobacteria; Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11889109;
RA Kapral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasilev O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fomstein M., Kyripides N., Overbeek R.,
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
nucleatum strain ATCC 25586";
RL J. Bacteriol. 184:2005-2018(2002).
CC -1- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +
phosphate + L-glutamate + carbamoyl phosphate.
CC -1- COPACTOR: Binds three manganese ions (By similarity).
CC -1- PATHWAY: Arginine biosynthesis.
CC -1- PATHWAY: Pyrimidine biosynthesis; first step.
CC -1- SUBUNIT: Composed of two chains; the small (or glutamine) chain
promotes the hydrolysis of glutamine to ammonia, which is used by
the large (or ammonia) chain to synthesize carbamoyl phosphate (By
similarity).
CC -1- SIMILARITY: BELONGS TO THE CARB FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE010554; AAL94625.1; ALT_INIT.
DR InterPro; IPR005483; CPhase_L.
DR InterPro; IPR005479; CPhase_L_D2.
DR InterPro; IPR005480; CPhase_L_D3.


```

DR InterPro; IPR005481; CPase_L_N.
DR InterPro; IPR004362; MGS like.
DR Pfam; PF00289; CPase_L_Chain; 2.
DR Pfam; PF02786; CPase_L_D2; 2.
DR Pfam; PF02787; CPase_L_D3; 1.
DR Pfam; PF02142; MGS; 1.
DR PRINTS; PR00098; CPASE.
DR PROSITE; PS00866; CPASE 1; 2.
DR PROSITE; PS00867; CPASE 2; 2.
KW Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;
KW ATP-binding; Manganese; Complete proteome.
FT DOMAIN 1 401 CARBOXYPHOSPHATE SYNTHETIC DOMAIN.
FT DOMAIN 402 546 OLIGOMERIZATION DOMAIN.
FT DOMAIN 547 929 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.
FT DOMAIN 930 1058 ALLOSTERIC DOMAIN.
FT REPEAT 1 546
FT REPEAT 547 1058
FT NP_BIND 153 210 ATP (POTENTIAL).
FT NP_BIND 302 352 ATP (POTENTIAL).
FT METAL 284 284 MANGANESE 1 (BY SIMILARITY).
FT METAL 298 298 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 300 300 MANGANESE 2 (BY SIMILARITY).
FT METAL 820 820 MANGANESE 3 (BY SIMILARITY).
FT METAL 832 832 MANGANESE 3 (BY SIMILARITY).
SQ SEQUENCE 1058 AA; 117451 MW; ED7037AF77C1E39F CRC64;

Query Match 65.5%; Score 36; DB 1; Length 1058;
Best Local Similarity 60.0%; Pred. No. 38;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPGMSYS 11
DB 190 EIVPGLNYS 199

RESULT 8
RL20 MYCPU STANDARD; PRT; 116 AA.
AC Q98QV0;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L20.
GN RPLT OR MYPU 2610.
OS Mycoplasma pulmonis.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2107;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=UAB CTIP;
RX MEDLINE=21267165; PubMed=11353084;
RA Chambaud I., Heilig R., Ferris S., Barbe V., Samson D., Galissien F.,
RA Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P.C.,
RA Blanchard A.;
RT "The complete genome sequence of the murine respiratory pathogen
RT Mycoplasma pulmonis."
RL Nucleic Acids Res. 29:2145-2153(2001).
CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA AND IS
CC NECESSARY TO THE IN VITRO ASSEMBLY PROCESS OF THE 50S RIBOSOMAL
CC SUBUNIT; IT IS NOT INVOLVED IN THE PROTEIN SYNTHESIZING FUNCTIONS
CC OF THAT SUBUNIT (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE L20P FAMILY OF RIBOSOMAL PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AL445563; CAC13434.1; -.
CC Mypulist; MYPU_2610; -.

```

```

DR InterPro; IPR001081; Ribosomal L20.
DR Pfam; PF00453; Ribosomal L20; 1.
DR PRINTS; PR00062; RIBOSOMAL20.
DR ProDom; PD002389; Ribosomal L20; 1.
DR TIGRfam; TIGR01032; rplT bact; 1.
DR PROSITE; PS00937; RIBOSOMAL_L20; 1.
KW Ribosomal protein; rRNA-binding; Complete proteome.
SQ SEQUENCE 116 AA; 13565 MW; C59C748901B18F14 CRC64;

Query Match 63.6%; Score 35; DB 1; Length 116;
Best Local Similarity 77.8%; Pred. No. 6.1;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPVGMYS 11
DB 68 VRPLGMSYS 76

RESULT 9
CXAG6_RAT STANDARD; PRT; 286 AA.
AC P28233;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gap junction alpha-6 protein (Connexin 33) (Cx33).
GN GJA6 OR CXN-33.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=92112940; PubMed=1370487;
RA Haefliger J.-A., Bruzone R., Jenkins N.A., Gilbert D.J.,
RA Copeland N.G., Paul D.L.;
RT "Four novel members of the connexin family of gap junction proteins.
RT Molecular cloning, expression, and chromosome mapping."
RL J. Biol. Chem. 267:2057-2064(1992).
CC -1- FUNCTION: ONE GAP JUNCTION CHANNELS, THE CONNEXONS, THROUGH WHICH
CC MATERIALS OF LOW MW DIFFUSE FROM ONE CELL TO A NEIGHBORING CELL.
CC -1- SUBUNIT: A CONNEXON IS COMPOSED OF A HEXAMER OF CONNEXINS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN TESTIS.
CC -1- SIMILARITY: BELONGS TO THE CONNEXIN FAMILY. ALPHA-TYPE (GROUP II)
CC SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; M76534; AAA40998.1; -.
CC PIR; C42053; C42053.
DR InterPro; IPR000500; Connexin.
DR Pfam; PF00029; connexin; 1.
DR PRINTS; PR00206; CONNEXIN.
DR SMART; SMO0037; CNX; 1.
DR PROSITE; PS00407; CONNEXINS 1; 1.
DR PROSITE; PS00408; CONNEXINS 2; 1.
KW Gap junction; Transmembrane.
FT DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 21 41 POTENTIAL..
FT DOMAIN 42 76 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 77 97 POTENTIAL.
FT DOMAIN 98 163 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 164 184 POTENTIAL.
FT DOMAIN 185 208 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 209 229 POTENTIAL..

```

FT DOMAIN 230 286 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 286 AA; 32860 MW; A585266ACA2ACCF2 CRC64;

Query Match 63.6%; Score 35; DB 1; Length 286;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 BEVPGVMS 9
Db 274 DQVPGVGLS 282

RESULT 10
PLSB_HUMAN STANDARD; PRT; 828 AA.
AC Q9HCL2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycerol-3-phosphate acyltransferase, mitochondrial precursor
DE (EC 2.3.1.15) (GPAT).
GN GPAT OR KIAA1560.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ashwell R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 167-828 FROM N.A.
RX TISSUE=Brain;
RX MEDLINE=20450683; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.
RT XVIII. The complete sequences of 100 new cDNA clones from brain which
RT code for large proteins in vitro.";
RL DNA Res. 7:273-281(2000).
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + sn-glycerol 3-phosphate = CoA + 1-
CC acyl-sn-glycerol 3-phosphate.
CC -1- PATHWAY: FIRST STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS. IT MAY
CC ALSO FUNCTION IN THE REGULATION OF MEMBRANE BIOGENESIS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC outer membrane (By similarity).
CC -1- SIMILARITY: BELONGS TO THE GPAT / DAPAT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL391986; -; NOT ANNOTATED CDS.
CC EMBL; AB046780; BAB13386.1; -;
CC MIM; 602395; -;
CC InterPro; IPR002123; Acyltransferase.
CC Pfam; PF01553; Acyltransferase; 1.
CC Phospholipid biosynthesis; transferase; Acyltransferase;
CC Transmembrane; Mitochondrion; Trans peptide.
CC TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
CC CHAIN ? 828 MITOCHONDRIAL INTERMEMBRANE (POTENTIAL).
CC DOMAIN ? 471 MITOCHONDRIAL INTERMEMBRANE (POTENTIAL).
CC TRANSNEM 472 494 POTENTIAL.
CC DOMAIN 495 574 CYTOPLASMIC (POTENTIAL).
CC TRANSNEM 575 593 POTENTIAL.
CC DOMAIN 594 828 MITOCHONDRIAL INTERMEMBRANE (POTENTIAL).
CC SEQUENCE 828 AA; 93836 MW; E5CD321A23D0B65B CRC64;

Query Match 63.6%; Score 35; DB 1; Length 828;
Best Local Similarity 62.5%; Pred. No. 46;

Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VVPVMSY 10
Db 349 IIPVGISY 356

RESULT 11
ZEP1_HUMAN STANDARD; PRT; 2717 AA.
AC F15622;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Zinc finger protein 40 (Human immunodeficiency virus type I enhancer-
DE binding protein 1) (HIV-Ep1) (Major histocompatibility complex binding
DE protein 1) (MBP-1) (Positive regulatory domain II binding factor 1)
DE (PRDII-BF1).
GN HIVEP1 OR ZNF40.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90169514; PubMed=2106471;
RA Fan C.M., Maniatis T.;
RT "A DNA-binding protein containing two widely separated zinc finger
RT motifs that recognize the same DNA sequence.";
RL Genes Dev. 4:29-42(1990).
RN [2]
RP STRUCTURE BY NMR OF 2113-2142.
RX MEDLINE=91064333; PubMed=2248949;
RA Omichinski J.G., Clore G.M., Appella E., Sakaguchi K.,
RA Gronenborn A.M.;
RT "High-resolution three-dimensional structure of a single zinc finger
RT from a human enhancer binding protein in solution.";
RL Biochemistry 29:9324-9334(1990).
RN [3]
RP STRUCTURE BY NMR OF 2087-2142.
RX MEDLINE=92232684; PubMed=1567844;
RA Omichinski J.G., Clore G.M., Robien M., Sakaguchi K., Appella E.,
RA Gronenborn A.M.;
RT "High-resolution solution structure of the double Cys2His2 zinc
RT finger from the human enhancer binding protein MBP-1.";
RL Biochemistry 31:3907-3917(1992).
CC -1- FUNCTION: THIS PROTEIN SPECIFICALLY BINDS TO THE DNA SEQUENCE
CC 5'-GGGACTTCC-3', WHICH IS FOUND IN THE ENHANCER ELEMENTS OF
CC NUMEROUS VIRAL PROMOTERS SUCH AS THOSE OF SV40, CMV, OR HIV1.
CC IN ADDITION, RELATED SEQUENCES ARE FOUND IN THE ENHANCER ELEMENTS
CC OF A NUMBER OF CELLULAR PROMOTERS, INCLUDING THOSE OF THE CLASS I
CC MHC, INTERLEUKIN-2 RECEPTOR, AND INTERFERON-BETA GENES. IT MAY ACT
CC IN T-CELL ACTIVATION.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- INDUCTION: BY MITOGEN AND PHORBOL ESTER.
CC -1- DOMAIN: CONTAINS TWO SETS OF 2 ZINC-FINGERS, WHICH ARE WIDELY
CC SEPARATED AND RECOGNIZE THE SAME DNA SEQUENCE. THERE IS A FIFTH
CC ZINC-FINGER IN-BETWEEN
CC -1- SIMILARITY: STRONG, TO HIVEP2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X51435; CAA35798.1; -;
CC PIR; A34203; A34203.
CC PDB; 3ZNF; 15-JAN-92.
CC PDB; 4ZNF; 15-JAN-92.
CC PDB; 1BBO; 31-OCT-93.

DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE G1/S-specific cyclin D2 (Vin-1 proto-oncogene).
 GN CCND2 OR VIN-1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93275661; PubMed=8502486;
 RA Hanna Z., Jankowski M., Tremblay P., Jiang X.M., Milatovich A.,
 RA Francke U., Jolicœur P.;
 RT "The Vin-1 gene, identified by provirus insertional mutagenesis, is
 RT the cyclin D2";
 RL Oncogene 8:1661-1666(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93275661; PubMed=8502486;
 RA Hosokawa Y., Onga T., Nakashima K.;
 RA "Induction of D2 and D3 cyclin-encoding genes during promotion of the
 RT G1/S transition by prolactin in rat Nb2 cells";
 RL Gene 147:249-252(1994).
 CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
 CC (START) TRANSITION.
 CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
 CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
 CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
 CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L09752; AAA41010.1; -
 CC EMBL; D16308; BAA03815.1; -
 CC InterPro; IPR004366; Cyclin.
 CC InterPro; IPR004367; Cyclin_Cterm.
 CC Pfam; PF00134; cyclin; 1.
 CC SMART; SM00385; CYCLIN; 1.
 CC PROSITE; PS00292; CYCLIN; 1.
 CC Cyclin; Cell cycle; Cell division; Multigene family; Proto-oncogene.
 KW CONFLICT 68 68 E -> G (IN REF. 2).
 FT CONFLICT 104 104 C -> V (IN REF. 2).
 FT CONFLICT 232 232 T -> A (IN REF. 2).
 SQ SEQUENCE 288 AA; 32826 MW; 4B522BF4E9835FC1 CRC64;

 Query Match 61.8%; Score 34; DB 1; Length 288;
 Best Local Similarity 60.0%; Pred. No. 25;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

 QY 1 EEVVPVGMYS 10
 ||| | : | : |
 Db 73 EEVFPPLANNY 82

 RESULT 15
 CGD2_HUMAN
 ID CGD2_HUMAN STANDARD; PRT; 289 AA.
 AC P30279; Q13955;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE G1/S-specific cyclin D2.
 GN CCND2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92347851; PubMed=1386336;
 RA Xiong Y., Menninger J., Beach D., Ward D.C.;
 RT "Molecular cloning and chromosomal mapping of CCND genes encoding
 RT human D-type cyclins";
 RL Genomics 13:575-584(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92347851; PubMed=1386336;
 RA Palmero I., Holder A., Sinclair A.J., Dickson C., Peters G.;
 RA "Cyclins D1 and D2 are differentially expressed in human B-lymphoid
 RT cell lines";
 RL Oncogene 8:1049-1054(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92347851; PubMed=1386336;
 RA Miyajima N.;
 RA Submitted (MAR-1993) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92347850; PubMed=1386335;
 RA Inaba T., Matsushima H., Valentine M., Roussel M.F., Sherr C.J.,
 RA Look A.T.;
 RA "Genomic organization, chromosomal localization, and independent
 RT expression of human cyclin D genes";
 RL Genomics 13:565-574(1992).
 CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
 CC (START) TRANSITION.
 CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
 CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
 CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
 CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M90813; AAA51926.1; -
 CC EMBL; X68452; CAA48493.1; -
 CC EMBL; D13639; BAA02802.1; -
 CC EMBL; BC010958; AAH10958.1; -
 CC EMBL; M88083; AAA51928.1; -
 CC EMBL; M88080; AAA51928.1; JOINED.
 CC EMBL; M88081; AAA51928.1; JOINED.
 CC EMBL; M88082; AAA51928.1; JOINED.
 CC PIR; A42822; A42822.
 CC PIR; S26580; S26580.
 CC Genew; HGNC:1583; CCND2.
 CC MIM; 123833; -
 CC InterPro; IPR004366; Cyclin.
 CC InterPro; IPR004367; Cyclin_Cterm.
 CC Pfam; PF00134; cyclin; 1.
 CC Pfam; PF02984; cyclin; 1.
 CC SMART; SM00385; CYCLIN; 1.
 CC PROSITE; PS00292; CYCLIN; 1.
 CC Cyclin; Cell cycle; Cell division; Multigene family.
 KW CONFLICT 166 167 KL -> NV (IN REF. 5).
 FT CONFLICT 224 224 T -> H (IN REF. 5).
 SQ SEQUENCE 289 AA; 33067 MW; B4E5FEF476D76D90 CRC64;

 Query Match 61.8%; Score 34; DB 1; Length 289;
 Best Local Similarity 60.0%; Pred. No. 25;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPVGMYSY 10
|||:|:
Db 74 EEVFPPLANNY 83

Search completed: June 4, 2003, 13:11:47
Job time : 7.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:13:35 ; Search time 14.25 Seconds
(without alignments)
79.694 Million cell updates/sec

Title: AUDET-909-4

Perfect score: 55

Sequence: 1 eevvpvqmsys 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:**

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	65.5	3472	9	US-10-027-806-4
2	36	65.5	3472	9	US-10-034-623-4
3	36	65.5	3472	9	US-10-027-801-4
4	35	63.6	332	9	US-09-738-626-6675
5	35	63.6	828	9	US-09-935-290-2
6	34	61.8	254	10	US-09-778-927A-53
7	34	61.8	289	9	US-10-024-066-2
8	34	61.8	289	10	US-10-024-066-4
9	34	61.8	289	10	US-09-919-497-54
10	34	61.8	691	9	US-10-101-921-4
11	34	61.8	591	10	US-09-925-731-2
12	34	61.8	774	9	US-10-270-333-111
13	33	60.0	10	10	US-09-947-387-66
14	33	60.0	426	9	US-10-214-766-43
15	33	60.0	478	9	US-09-924-340-108
16	33	60.0	478	9	US-09-992-600A-108
17	33	60.0	478	9	US-09-746-783-184
18	33	60.0	478	9	US-10-000-489-108
19	33	60.0	478	9	US-10-000-986-108

20	33	60.0	622	9	US-09-738-626-4919	Sequence 4919, Ap
21	33	60.0	702	9	US-10-280-403-2	Sequence 2, Appli
22	33	60.0	702	10	US-09-907-479-2	Sequence 2, Appli
23	32	58.2	108	10	US-09-817-182-2	Sequence 2, Appli
24	32	58.2	295	10	US-09-925-300-1061	Sequence 1061, Ap
25	32	58.2	529	10	US-09-923-304-4	Sequence 4, Appli
26	32	58.2	679	10	US-09-815-242-5658	Sequence 5658, Ap
27	32	58.2	681	10	US-09-815-242-12270	Sequence 12270, A
28	32	58.2	1377	9	US-10-101-464A-73	Sequence 73, Appl
29	32	58.2	1377	10	US-09-815-242-10384	Sequence 10384, A
30	32	58.2	2799	9	US-10-151-736-4	Sequence 4, Appli
31	32	58.2	3256	10	US-09-919-172-98	Sequence 98, Appl
32	31	56.4	53	9	US-10-092-154-878	Sequence 878, App
33	31	56.4	53	10	US-09-764-847-878	Sequence 878, App
34	31	56.4	72	10	US-09-864-761-34309	Sequence 34309, A
35	31	56.4	299	10	US-09-815-242-10697	Sequence 10697, A
36	31	56.4	335	10	US-09-925-300-1575	Sequence 1575, Ap
37	31	56.4	395	10	US-09-815-242-11724	Sequence 11724, A
38	31	56.4	422	9	US-10-169-048-62	Sequence 62, Appl
39	31	56.4	426	10	US-09-815-242-5383	Sequence 5383, Ap
40	31	56.4	429	10	US-09-815-242-12502	Sequence 12502, A
41	31	56.4	477	10	US-09-815-242-13881	Sequence 13881, A
42	31	56.4	546	9	US-10-153-668-356	Sequence 356, App
43	31	56.4	563	9	US-10-153-668-236	Sequence 236, App
44	31	56.4	564	10	US-09-764-864-1245	Sequence 1245, Ap
45	31	56.4	573	9	US-10-153-668-234	Sequence 234, App

ALIGNMENTS

RESULT 1

US-10-027-806-4

; Sequence 4, Application US/10027806

; Patent No. US20020160476A1

; GENERAL INFORMATION:

; APPLICANT: Swanson, Ronald V.

; APPLICANT: Feldman, Robert A.

; APPLICANT: Schleper, Christa

; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM

; FILE REFERENCE: DCORP.002A

; CURRENT APPLICATION NUMBER: US/10/027,806

; PRIOR FILING DATE: 2001-12-21

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020

; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29

; NUMBER OF SEQ ID NOS: 123

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 4

; LENGTH: 3472

; TYPE: PRT

; ORGANISM: Cenarchaeum symbiosum

US-10-027-806-4

Query Match 65.5%; Score 36; DB 9; Length 3472;

Best Local Similarity 54.5%; Pred. No. 5.8e+02;

Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEVVPVQMSYS 11

|:|:| |:|

Db 2294 EDVIPRGISFS 2304

RESULT 2

US-10-034-623-4

; Sequence 4, Application US/10034623

; Publication No. US20020198365A1

; GENERAL INFORMATION:

; APPLICANT: Swanson, Ronald V.

; APPLICANT: Feldman, Robert A.

; APPLICANT: Schleper, Christa

; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM

; FILE REFERENCE: DCORP.002A

; CURRENT APPLICATION NUMBER: US/10/034,623

; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: 60/102,294
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 3472
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-10-034-623-4

Query Match 65.5%; Score 36; DB 9; Length 3472;
Best Local Similarity 54.5%; Pred. No. 5.8e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVVVPVGMYS 11
|:|:| |:|:|
Db 2294 EDVIPRGISFS 2304

RESULT 3

US-10-027-801-4
; Sequence 4, Application US/10027801
; Publication No. US20030054364A1
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCRP.002A
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US/10/027,801
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 3472
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-10-027-801-4

Query Match 65.5%; Score 36; DB 9; Length 3472;
Best Local Similarity 54.5%; Pred. No. 5.8e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVVVPVGMYS 11
|:|:| |:|:|
Db 2294 EDVIPRGISFS 2304

RESULT 4

US-09-738-626-6675
; Sequence 6675, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484

; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6675
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6675

Query Match 63.6%; Score 35; DB 9; Length 332;
Best Local Similarity 75.0%; Pred. No. 71;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VVPVGMYS 10
|:|:| |:|:|
Db 118 VLPVGMAY 125

RESULT 5

US-09-935-290-2
; Sequence 2, Application US/09935290
; Publication No. US20030044948A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 56919, A NOVEL HUMAN ACYLTRANSFERASE AND USES THEREOF
; FILE REFERENCE: WNI-186
; CURRENT APPLICATION NUMBER: US/09/935,290
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: 60/226,509
; PRIOR FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 2
; LENGTH: 828
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-935-290-2

Query Match 63.6%; Score 35; DB 9; Length 828;
Best Local Similarity 62.5%; Pred. No. 1.9e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VVPVGMYS 10
|:|:| |:|:|
Db 349 IIPVGISY 356

RESULT 6

US-09-778-927A-53
; Sequence 53, Application US/09778927A
; Patent No. US20020068342A1
; GENERAL INFORMATION:
; APPLICANT: KHOSRAVI, Rami et al.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID AND AMINO ACID SEQUENCES AND NOVEL
; TITLE OF INVENTION: VARIANTS OF ALTERNATIVE SPLICING
; FILE REFERENCE: 2786-0160P
; CURRENT APPLICATION NUMBER: US/09/778,927A
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: IL 134453
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: IL135341
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 53
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

; NAME/KEY: misc_feature
; LOCATION: (1)..(254)
; OTHER INFORMATION: Xaa = any amino acid, unknown, or other
US-09-778-927A-53

Query Match 61.8%; Score 34; DB 10; Length 254;
Best Local Similarity 60.0%; Pred. No. 83;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 BEVVPVGMYSY 10
||| |.:.|
Db 74 BEVFPPLANNY 83

RESULT 7

US-10-024-066-2
; Sequence 2, Application US/10024066
; Patent No. US20020166134A1

; GENERAL INFORMATION:

; APPLICANT: Field, Loren J.

; APPLICANT: Pasumarthi, Kishore Babu S.

; TITLE OF INVENTION: CARDIOMYOCYTES WITH ENHANCED PROLIFERATIVE POTENTIAL,

; TITLE OF INVENTION: AND METHODS FOR PREPARING AND USING SAME

; FILE REFERENCE: 7037-450

; CURRENT APPLICATION NUMBER: US/10/024.066

; PRIOR FILING DATE: 2001-12-18

; PRIOR FILING DATE: 2001-12-18

; PRIOR FILING DATE: 1999-06-18

; PRIOR FILING DATE: PCT/US00/16827

; PRIOR FILING DATE: 2000-06-19

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 289

; TYPE: PRT

; ORGANISM: Mus musculus

US-10-024-066-2

Query Match 61.8%; Score 34; DB 9; Length 289;
Best Local Similarity 60.0%; Pred. No. 95;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 BEVVPVGMYSY 10
||| |.:.|
Db 73 BEVFPPLANNY 82

RESULT 8

US-10-024-066-4

; Sequence 4, Application US/10024066

; Patent No. US20020166134A1

; GENERAL INFORMATION:

; APPLICANT: Field, Loren J.

; APPLICANT: Pasumarthi, Kishore Babu S.

; TITLE OF INVENTION: CARDIOMYOCYTES WITH ENHANCED PROLIFERATIVE POTENTIAL,

; TITLE OF INVENTION: AND METHODS FOR PREPARING AND USING SAME

; FILE REFERENCE: 7037-450

; CURRENT APPLICATION NUMBER: US/10/024.066

; PRIOR FILING DATE: 2001-12-18

; PRIOR FILING DATE: 2001-12-18

; PRIOR FILING DATE: 1999-06-18

; PRIOR FILING DATE: PCT/US00/16827

; PRIOR FILING DATE: 2000-06-19

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 289

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-024-066-4

Query Match 61.8%; Score 34; DB 9; Length 289;
Best Local Similarity 60.0%; Pred. No. 95;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 BEVVPVGMYSY 10
||| |.:.|
Db 74 BEVFPPLANNY 83

RESULT 9

US-09-919-497-54

; Sequence 54, Application US/09919497

; Patent No. US20020106662A1

; GENERAL INFORMATION:

; APPLICANT: Mutter, George L.

; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER

; FILE REFERENCE: B0801/7225

; CURRENT APPLICATION NUMBER: US/09/919.497

; PRIOR FILING DATE: 2001-07-31

; PRIOR APPLICATION NUMBER: US 60/221,735

; PRIOR FILING DATE: 2000-07-31

; NUMBER OF SEQ ID NOS: 100

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 54

; LENGTH: 289

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-919-497-54

Query Match 61.8%; Score 34; DB 10; Length 289;
Best Local Similarity 60.0%; Pred. No. 95;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 BEVVPVGMYSY 10
||| |.:.|
Db 74 BEVFPPLANNY 83

RESULT 10

US-10-101-921-4

; Sequence 4, Application US/10101921

; Publication No. US20030022199A1

; GENERAL INFORMATION:

; APPLICANT: Nezu, Jun-ichi

; APPLICANT: Ose, Asuka

; APPLICANT: Tsuji, Akira

; TITLE OF INVENTION: TRANSPORTER GENES OATP-B, C, D, AND E

; FILE REFERENCE: 06501-104US1

; CURRENT APPLICATION NUMBER: US/10/101.921

; PRIOR FILING DATE: 2002-06-28

; PRIOR APPLICATION NUMBER: PCT/JP00/06416

; PRIOR FILING DATE: 2000-09-20

; PRIOR APPLICATION NUMBER: JP 11/267835

; PRIOR FILING DATE: 1999-09-21

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 691

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-101-921-4

Query Match 61.8%; Score 34; DB 9; Length 691;
Best Local Similarity 62.5%; Pred. No. 2.4e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 VVPVGMYSY 10
:|:|:|:|
Db 188 IVPLGLSY 195

RESULT 11

US-09-925-731-2

; Sequence 2, Application US/09925731

; Patent No. US20020090622A1

; GENERAL INFORMATION:
; APPLICANT: ADEKUN, ANTHONI MONISOLA
; APPLICANT: AMBROSE, HELEN JEAN
; APPLICANT: CRESSWELL, CARL JOHN
; APPLICANT: DUDLEY, ADAM JESTON
; TITLE OF INVENTION: CHEMICAL COMPOUNDS
; FILE REFERENCE: DUB/009901/0282795
; CURRENT APPLICATION NUMBER: US/09/925,731
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/226,909
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-731-2

Query Match 61.8%; Score 34; DB 10; Length 691;
Best Local Similarity 62.5%; Pred. No. 2.4e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VVPVGMYS 10
Db 188 IVPLGLSY 195
:|:|:|:|

RESULT 12
US-10-270-333-111
; Sequence 111, Application US/10270333
; Publication No. US20030092124A1
; GENERAL INFORMATION:
; APPLICANT: Cravchik, Anibal
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF AS INSECTICIDAL TARGETS
; FILE REFERENCE: CU000733CON
; CURRENT APPLICATION NUMBER: US/10/270,333
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/168,677
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: 60/175,691
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/191,638
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 198
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 111
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Drosophila
US-10-270-333-111

Query Match 61.8%; Score 34; DB 9; Length 774;
Best Local Similarity 63.8%; Pred. No. 2.8e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EEVVPVGMYS 11
Db 753 EESVPVGLTTS 763
||| ||||:|

RESULT 13
US-09-947-387-66
; Sequence 66, Application US/09947387
; Patent No. US20020150885A1
; GENERAL INFORMATION:
; APPLICANT: Weber, Eckard
; APPLICANT: Cai, Sui Xiong
; APPLICANT: Keana, John F.W.
; APPLICANT: Drewe, John A.
; APPLICANT: Zhang, Han-Zhong

; TITLE OF INVENTION: No. US20020150885A1 Fluorogenic or Fluorescent Reporter Molecule
; TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence
; TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the
; TITLE OF INVENTION: Use Thereof
; FILE REFERENCE: 1735.0290005
; CURRENT APPLICATION NUMBER: US/09/947,387
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/061,582
; PRIOR FILING DATE: 1997-10-10
; PRIOR APPLICATION NUMBER: US 60/145,746
; PRIOR FILING DATE: 1998-03-03
; PRIOR APPLICATION NUMBER: US 09/168,888
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 66
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-947-387-66

Query Match 60.0%; Score 33; DB 10; Length 10;
Best Local Similarity 50.0%; Pred. No. 3.9;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EEVVPVGMYS 10
Db 1 DDIVPCMSY 10
:|:|:|

RESULT 14
US-10-214-766-43
; Sequence 43, Application US/10214766
; Publication No. US20030084473A1
; GENERAL INFORMATION:
; APPLICANT: Gocal, Greg
; TITLE OF INVENTION: NON-TRANSGENIC HERBICIDE RESISTANT PLANTS
; FILE REFERENCE: CA1138
; CURRENT APPLICATION NUMBER: US/10/214,766
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,734
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 43
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-214-766-43

Query Match 60.0%; Score 33; DB 9; Length 426;
Best Local Similarity 60.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPVGMYS 10
Db 223 EFVIPAGQSY 232
||:|:|

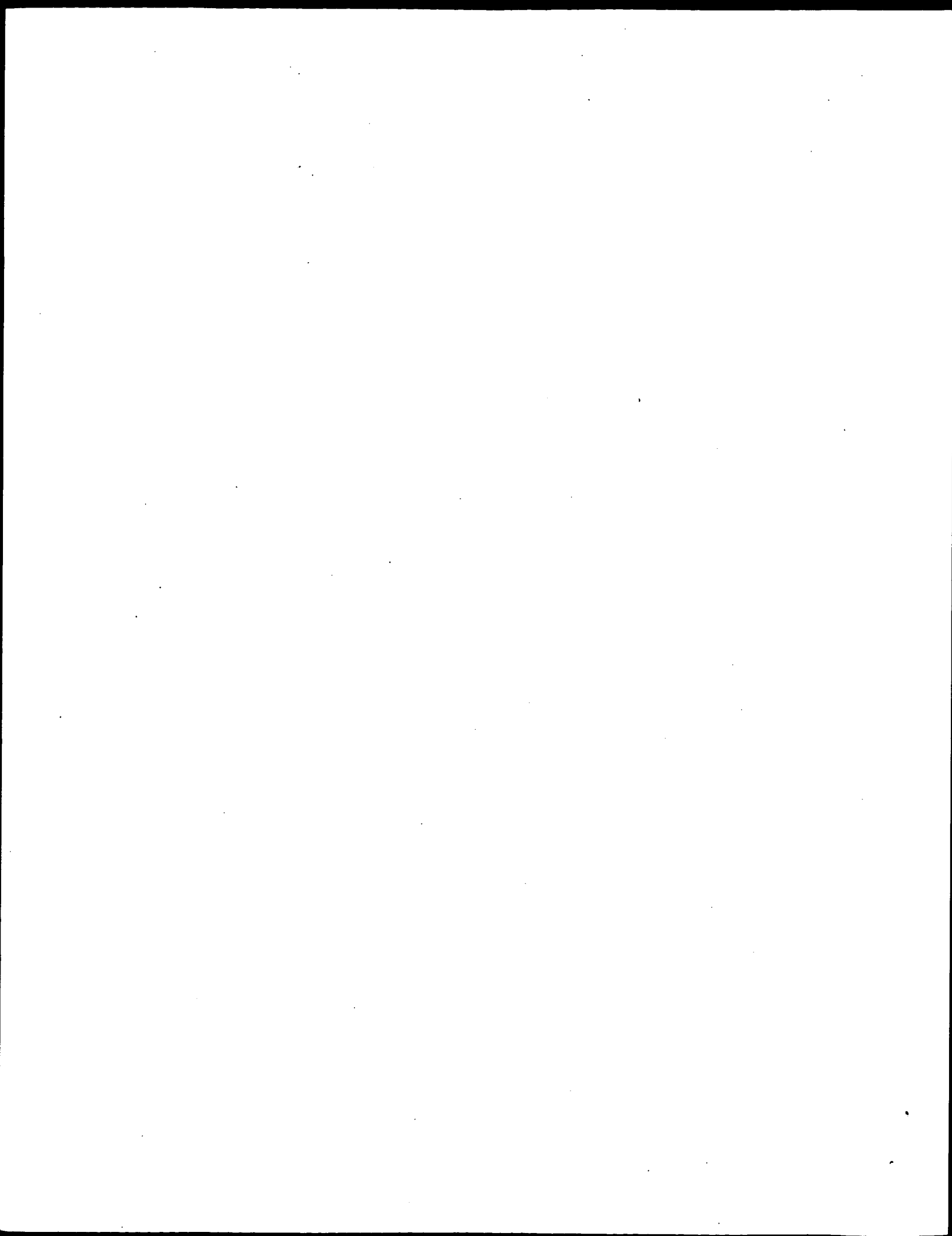
RESULT 15
US-09-924-340-108
; Sequence 108, Application US/09924340
; Publication No. US20030027248A1
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US2.REG
; CURRENT APPLICATION NUMBER: US/09/924,340
; CURRENT FILING DATE: 2001-08-06

; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 108
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-924-340-108

Query Match 60.0%; Score 33; DB 9; Length 478;
Best Local Similarity 60.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPVGM SYS 11
|||
Db 239 EVAPAGASYN 248

Search completed: June 4, 2003, 13:30:39
Job time : 14.25 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 4, 2003, 13:09:19 ; Search time 9.75 Seconds
(without alignments)
33.195 Million cell updates/sec

Title: AUDET-909-4
Perfect score: 55
Sequence: 1 eevpvgmsys 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	63.6	45	2	US-08-637-759B-236
2	35	63.6	45	3	US-08-871-355A-236
3	35	63.6	45	4	US-09-201-945-236
4	34	61.8	102	2	US-08-580-988A-23
5	34	61.8	152	2	US-08-460-694-4
6	34	61.8	152	3	US-08-460-744-4
7	34	61.8	152	3	US-07-667-711B-4
8	34	61.8	173	1	US-08-193-977-7
9	34	61.8	189	2	US-08-464-517-21
10	34	61.8	189	2	US-08-246-361A-21
11	34	61.8	189	3	US-08-463-772-21
12	34	61.8	189	5	PCT-US93-05000-21
13	34	61.8	236	2	US-08-464-517-22
14	34	61.8	236	2	US-08-246-361A-22
15	34	61.8	236	3	US-08-463-772-22
16	34	61.8	236	5	PCT-US93-05000-22
17	34	61.8	280	2	US-08-464-517-6
18	34	61.8	280	3	US-08-463-772-6
19	34	61.8	289	2	US-08-246-361A-4
20	34	61.8	289	5	PCT-US93-05000-4
21	34	61.8	291	5	PCT-US93-05000-6
22	34	61.8	292	2	US-08-464-517-23
23	34	61.8	292	2	US-08-246-361A-6
24	34	61.8	292	2	US-08-246-361A-23
25	34	61.8	292	3	US-08-463-772-23
26	34	61.8	292	5	PCT-US93-05000-23
27	34	61.8	295	1	US-07-947-120-8

28	34	61.8	295	1	US-08-472-893A-8	Sequence 8, Appli
29	34	61.8	295	2	US-08-460-694-2	Sequence 2, Appli
30	34	61.8	295	2	US-08-464-517-19	Sequence 19, Appli
31	34	61.8	295	2	US-08-464-517-20	Sequence 20, Appli
32	34	61.8	295	2	US-08-246-361A-19	Sequence 19, Appli
33	34	61.8	295	2	US-08-246-361A-20	Sequence 20, Appli
34	34	61.8	295	3	US-08-463-772-19	Sequence 19, Appli
35	34	61.8	295	3	US-08-463-772-20	Sequence 20, Appli
36	34	61.8	295	3	US-08-460-744-2	Sequence 2, Appli
37	34	61.8	295	3	US-07-667-711B-2	Sequence 2, Appli
38	34	61.8	295	3	US-08-947-492-8	Sequence 8, Appli
39	34	61.8	295	5	PCT-US93-05000-2	Sequence 2, Appli
40	34	61.8	295	5	PCT-US93-05000-19	Sequence 19, Appli
41	34	61.8	295	5	PCT-US93-05000-20	Sequence 20, Appli
42	34	61.8	309	2	US-08-464-517-4	Sequence 4, Appli
43	34	61.8	309	3	US-08-463-772-4	Sequence 4, Appli
44	34	61.8	618	2	US-08-770-761A-3	Sequence 3, Appli
45	34	61.8	647	2	US-08-770-761A-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-08-637-759B-236
; Sequence 236, Application US/08637759B
; Patent No. 5876931
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/637,759B
; FILING DATE: 03-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 236:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-637-759B-236

Query Match 63.6%; Score 35; DB 2; Length 45;
Best Local Similarity 60.0%; Pred. No. 3.6;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 EVVVPVGMYS 10
||: ||: ||
Db 1 EISPLGWSY 10

RESULT 2

US-08-871-355A-236
; Sequence 236, Application US/08871355A
; Patent No. 6015669
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/871,355A
; FILING DATE: 09-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101 CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 236:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-871-355A-236

Query Match 63.6%; Score 35; DB 3; Length 45;
Best Local Similarity 60.0%; Pred. No. 3.6;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 EVVVPVGMYS 10
||: ||: ||
Db 1 EISPLGWSY 10

RESULT 3

US-09-201-945-236
; Sequence 236, Application US/09201945
; Patent No. 6342215
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta

STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/201,945
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/637,759
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPMS 101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 236:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-201-945-236

Query Match 63.6%; Score 35; DB 4; Length 45;
Best Local Similarity 60.0%; Pred. No. 3.6;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 EVVVPVGMYS 10
||: ||: ||
Db 1 EISPLGWSY 10

RESULT 4

US-08-580-988A-23
; Sequence 23, Application US/08580988A
; Patent No. 5856161
; GENERAL INFORMATION:
; APPLICANT: Aggarwal et al.
; TITLE OF INVENTION: Tumor Necrosis Factor
; TITLE OF INVENTION: Receptor-I-Associated Protein Kinase And Methods
; TITLE OF INVENTION: For Its Use
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. Benjamin A. Adler
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/580,988A
; FILING DATE: January 3, 1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benjamin Aaron Adler, Ph.D., J.D.

REGISTRATION NUMBER: 35,423
REFERENCE/DOCKET NUMBER: D5721CIP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713-777-2321
TELEFAX: 713-777-6908
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 102 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: no
HYPOTHETICAL: no
ANTI-SENSE: no
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
US-08-580-988A-23

Query Match 61.8%; Score 34; DB 2; Length 102;
Best Local Similarity 60.0%; Pred. No. 14;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 BEVVPVGMYSY 10
Db 24 BEVFPPLAMNY 33

RESULT 5
US-08-460-694-4
Sequence 4, Application US/08460694
Patent No. 5858655
GENERAL INFORMATION:
APPLICANT: Arnold, Andrew
TITLE OF INVENTION: Pradi Cyclin and its cDNA
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 New York Avenue, N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,694
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: McConathy, Evelyn H.
REGISTRATION NUMBER: 35,279
REFERENCE/DOCKET NUMBER: 0609.4070002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 152 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-460-694-4

Query Match 61.8%; Score 34; DB 2; Length 152;
Best Local Similarity 60.0%; Pred. No. 23;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 BEVVPVGMYSY 10

Db 20 BEVFPPLAMNY 29

RESULT 6
US-08-460-744-4
Sequence 4, Application US/08460744
Patent No. 6107541
GENERAL INFORMATION:
APPLICANT: Arnold, Andrew
TITLE OF INVENTION: Pradi Cyclin and its cDNA
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 New York Avenue, N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,744
FILING DATE: 02-JUN-1995
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: McConathy, Evelyn H.
REGISTRATION NUMBER: 35,279
REFERENCE/DOCKET NUMBER: 0609.4070005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 152 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-460-744-4

Query Match 61.8%; Score 34; DB 3; Length 152;
Best Local Similarity 60.0%; Pred. No. 23;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 BEVVPVGMYSY 10
Db 20 BEVFPPLAMNY 29

RESULT 7
US-07-667-711B-4
Sequence 4, Application US/07667711B
Patent No. 6110700
GENERAL INFORMATION:
APPLICANT: ARNOLD, ANDREW
TITLE OF INVENTION: Pradi Cyclin and its cDNA
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVE., NW, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/667,711B
; FILING DATE: 11-MAR-1991
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MCPHAIL, DONALD R.
; REGISTRATION NUMBER: 35,811
; REFERENCE/DOCKET NUMBER: 0609.4070000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 152 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-07-667-711B-4

Query Match 61.8%; Score 34; DB 3; Length 152;
Best Local Similarity 60.0%; Pred. No. 23;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVVVPVGMYS 10
Db 20 EEVFFPLAMNY 29

RESULT 8
US-08-193-977-7
; Sequence 7, Application US/08193977
; Patent No. 5625031
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, KEVIN R.
; APPLICANT: COLEMAN, KEVIN G.
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF THE P33CDK2 AND
; TITLE OF INVENTION: P34CDC2 CELL CYCLE REGULATORY KINASES AND HUMAN
; TITLE OF INVENTION: PAPILLOMAVIRUS E7 ONCOPROTEIN
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/193,977
; FILING DATE: 08-FEB-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 5998-0016
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 173 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-193-977-7

Query Match 61.8%; Score 34; DB 1; Length 173;
Best Local Similarity 60.0%; Pred. No. 23;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVVVPVGMYS 10
Db 20 EEVFFPLAMNY 29

US-08-464-517-21
; Sequence 21, Application US/08464517
; Patent No. 5869640
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,517
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 189 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-464-517-21

Query Match 61.8%; Score 34; DB 2; Length 189;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVVVPVGMYS 10
Db 74 EEVFFPLAMNY 83

RESULT 10
US-08-246-361A-21
; Sequence 21, Application US/08246361A
; Patent No. 598582
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
```

```

Best Local Similarity 60.0%; Pred. No. 27;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVVVPVGMYS 10
Db 55 EEVFFPLAMNY 64

RESULT 9
US-08-464-517-21
; Sequence 21, Application US/08464517
; Patent No. 5869640
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,517
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 189 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-464-517-21

Query Match 61.8%; Score 34; DB 2; Length 189;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVVVPVGMYS 10
Db 74 EEVFFPLAMNY 83

RESULT 10
US-08-246-361A-21
; Sequence 21, Application US/08246361A
; Patent No. 598582
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
```


;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: LAHIVE & COCKFIELD
;; STREET: 60 State Street
;; CITY: Boston
;; STATE: MA
;; COUNTRY: USA
;; ZIP: 02109
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: ASCII(text)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/246,361A
;; FILING DATE: 19-MAY-1994
;; CLASSIFICATION: 435
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/963,308
;; FILING DATE: 16-OCT-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/888,178
;; FILING DATE: 26-MAY-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/701,514
;; FILING DATE: 16-MAY-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Matthew P. Vincent
;; REGISTRATION NUMBER: 36,709
;; REFERENCE/DOCKET NUMBER: MII-004C
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (617) 227-5941
;; TELEFAX: (617) 227-5941
;; INFORMATION FOR SEQ ID NO: 21:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 189 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;;
US-08-246-361A-21

Query Match 61.8%; Score 34; DB 2; Length 189;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVVVPVGMYSY 10
Db 74 EVVFLPMY 83

RESULT 11
US-08-463-772-21
;; Sequence 21, Application US/08463772
;; Patent No. 6066501
;; GENERAL INFORMATION:
;; APPLICANT: BEACH, David H.
;; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
;; NUMBER OF SEQUENCES: 50
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: LAHIVE & COCKFIELD
;; STREET: 60 State Street
;; CITY: Boston
;; STATE: MA
;; COUNTRY: USA
;; ZIP: 02109
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: ASCII(text)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/463,772
;; FILING DATE:

;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/963,308
;; FILING DATE: 16-OCT-1992
;; APPLICATION NUMBER: US 07/888,178
;; FILING DATE: 26-MAY-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/701,514
;; FILING DATE: 16-MAY-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Matthew P. Vincent
;; REGISTRATION NUMBER: 36,709
;; REFERENCE/DOCKET NUMBER: MII-004C
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (617) 227-7400
;; TELEFAX: (617) 227-5941
;; INFORMATION FOR SEQ ID NO: 21:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 189 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;;
US-08-463-772-21

Query Match 61.8%; Score 34; DB 3; Length 189;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVVVPVGMYSY 10
Db 74 EVVFLPMY 83

RESULT 12
PCT-US93-05000-21
;; Sequence 21, Application PC/TUS9305000
;; GENERAL INFORMATION:
;; APPLICANT: MITOTIX
;; TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto
;; NUMBER OF SEQUENCES: 42
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
;; STREET: Two Militia Drive
;; CITY: Lexington
;; STATE: Massachusetts
;; COUNTRY: US
;; ZIP: 02173
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US93/05000
;; FILING DATE: 19930525
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/07/888,178
;; FILING DATE: 26-MAY-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Granahan, Patricia
;; REGISTRATION NUMBER: 32,227
;; REFERENCE/DOCKET NUMBER: CSHL91-02A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 617-861-6240
;; TELEFAX: 616-861-9540
;; INFORMATION FOR SEQ ID NO: 21:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 189 amino acids
;; TYPE: AMINO ACID
;; TOPOLOGY: unknown
;; MOLECULE TYPE: protein

PCT-US93-05000-21

Query Match 61.8%; Score 34; DB 5; Length 189;
 Best Local Similarity 60.0%; Pred. No. 29;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EEVVPVGMYS 10
 |||||:|:
 Db 74 EEVFPLANNY 83

RESULT 13

US-08-464-517-22
 ; Sequence 22, Application US/08464517
 ; Patent No. 5869640
 ; GENERAL INFORMATION:
 ; APPLICANT: BEACH, David H.
 ; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
 ; NUMBER OF SEQUENCES: 50
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 60 State Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII(text)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/464,517
 ; FILING DATE: 16-OCT-1992
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/963,308
 ; FILING DATE: 16-OCT-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Matthew P. Vincent
 ; REGISTRATION NUMBER: 36,709
 ; REFERENCE/DOCKET NUMBER: MIL-004C
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 227-7400
 ; TELEFAX: (617) 227-5941
 ; INFORMATION FOR SEQ ID NO: 22:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 236 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-464-517-22

Query Match 61.8%; Score 34; DB 2; Length 236;
 Best Local Similarity 60.0%; Pred. No. 38;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EEVVPVGMYS 10
 |||||:|:
 Db 20 EEVFPLANNY 29

RESULT 14

US-08-246-361A-22
 ; Sequence 22, Application US/08246361A
 ; Patent No. 5998582
 ; GENERAL INFORMATION:

; APPLICANT: BEACH, David H.
 ; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
 ; NUMBER OF SEQUENCES: 50
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 60 State Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII(text)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/246,361A
 ; FILING DATE: 19-MAY-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/963,308
 ; FILING DATE: 16-OCT-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/888,178
 ; FILING DATE: 26-MAY-1992
 ; APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/701,514
 ; FILING DATE: 16-MAY-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Matthew P. Vincent
 ; REGISTRATION NUMBER: 36,709
 ; REFERENCE/DOCKET NUMBER: MIL-004C
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 227-7400
 ; TELEFAX: (617) 227-5941
 ; INFORMATION FOR SEQ ID NO: 22:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 236 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-246-361A-22

Query Match 61.8%; Score 34; DB 2; Length 236;
 Best Local Similarity 60.0%; Pred. No. 38;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EEVVPVGMYS 10
 |||||:|:
 Db 20 EEVFPLANNY 29

RESULT 15

US-08-463-772-22
 ; Sequence 22, Application US/08463772
 ; Patent No. 6066501
 ; GENERAL INFORMATION:
 ; APPLICANT: BEACH, David H.
 ; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
 ; NUMBER OF SEQUENCES: 50
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 60 State Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII(text)

;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/08/463.772
;/ FILING DATE:
;/ CLASSIFICATION: 435
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: US 07/963.308
;/ FILING DATE: 16-OCT-1992
;/ APPLICATION NUMBER: US 07/888.178
;/ FILING DATE: 26-MAY-1992
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: US 07/701.514
;/ FILING DATE: 16-MAY-1991
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: Matthew P. Vincent
;/ REGISTRATION NUMBER: 36,709
;/ REFERENCE/DOCKET NUMBER: MII-004C
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: (617) 227-7400
;/ TELEFAX: (617) 227-5941
;/ INFORMATION FOR SEQ ID NO: 22:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 236 amino acids
;/ TYPE: amino acid
;/ STRANDEDNESS: single
;/ TOPOLOGY: linear
;/ MOLECULE TYPE: peptide
;/ US-08-463-772-22

Query Match 61.8%; Score 34; DB 3; Length 236;
Best Local Similarity 60.0%; Pred. No. 38;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVFPVGMYSY 10
||| |:
Db 20 EEVFPPLANNY 29

Search completed: June 4, 2003, 13:14:59
Job time : 9.75 secs

